MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo •372

October 1976

ANALYSIS OF OCCLUDING CONTOUR

by

D. Marr

Summary.

Almost nothing can be deduced about a general 3-D surface given only its occluding contours in an image, yet contour information is easily and effectively used by us to infer the shape of a surface. Therefore, implicit in the perceptual analysis of occluding contour must lie various assumptions about the viewed surfaces. The assumptions that seem most natural are (a) that the distinction between convex and concave segments reflects real properties of the viewed surface; and (b) that contiguous portions of contour arise from contiguous parts of the viewed surface -- i.e., there are no invisible obscuring edges. It is proved that, for smooth surfaces, these assumptions are essentially equivalent to assuming that the viewed surface is a generalized cone. Methods are defined for finding the axis of such a cone, and for segmenting a surface constructed of several cones into its components, whose axes can then be found separately. These methods, together with the algorithms for implementing them devised by Vatan & Marr (1977), provide one link between an uninterpreted figure extracted from an image, and the 3-D representation theory of Marr & Nishihara (1977).

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology . Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643.

Introduction

When we look at the silhouettes in Picasso's work "Rites of Spring" (figure 1), we perceive them in terms of very particular 3-D shapes, some familiar, some less so. This is quite remarkable, because the silhouettes could in theory have been generated by an infinite variety of shapes which, from other viewpoints, have no discernable similarities to the shapes we perceive. One can perhaps attribute part of the phenomenon to a familiarity with the depicted shapes; but not all of it, because one can use the medium of a silhouette to convey a new shape, and because even with considerable effort it is difficult to imagine the more bizarre three-dimensional surfaces that could have given rise to the same silhouettes.

This phenomenon is of quite general importance for the analysis of an image. The boundary of a silhouette is simply one type of *occluding contour* (see e.g. Waltz 1975), and such contours are an artist's principal means of conveying information about shape. The paradox is that they apparently tell us more about shape than they should. For example, neighbouring points on such a contour could in general arise from widely separated points on the original surface, but our perceptual interpretation usually ignores this possibility.

This means that implicit in the way we interpret an occluding contour, there must lie some *a priori* assumptions that allow us to infer a shape from an outline. If a surface violates these assumptions, our analysis will be wrong, in the sense that the shape we assign to the contours will differ from the shape that actually caused them. An everyday example of this phenomenon is the shadowgraph, where the appropriate arrangement of one's hands can, to the surprise and delight of a child, produce the shadow of an apparently quite different shape, like a duck or a rabbit.

What assumptions is it reasonable to suppose that we make? I shall argue for these two; (a) that nearby points on a contour correspond to nearby points on the viewed surface; and (b) that the distinction between convexities and concavities in a contour reflects real properties of the surface, not an artifact of perspective.

Some surfaces seen from some viewpoints will satisfy these conditions, and some will not. Our first task is to understand what it is about a surface that makes it satisfy these assumptions, and the main result of the first part of the paper achieves this. Theorem 1 shows that, if the assumptions (a) and (b) hold for all distant vantage points such that the line of sight lies parallel to some fixed plane, then the viewed surface must be a *generalized cone*. (A generalized cone is the surface swept out by moving a cross-section of fixed shape but smoothly varying size, along an axis, as illustrated in figure 5).

This result is strong and surprising. It means that if one has a method for interpreting contours that relies on assumptions (a) and (b), then the method implicitly assumes that the viewed shape is a generalized cone. One can think of such a method as first throwing a generalized cone blanket round the viewed shape, and then describing the shape of the blanket. This in turn means that the representation of 3-D shape that is subsequently used can, without further loss of information, be based on generalized cones (like that of Marr & Nishihara 1977).

1. "Rites of spring" by Picasso. We immediately interpret the silhouettes in terms of particular 3-D surfaces, despite the paucity of information in the image. In order to do this, we must be bringing additional assumptions and constraints to bear on the analysis of these contours' shapes. This article enquires about the nature of this *a priori* information.

The result is of practical importance, because it can be used as a basis for methods that interpret occluding contours. This is the point of the later sections of the article, where we shall assume that the viewed object is a generalized cone, and ask how to discover its specifications (its axis and cross-section) using only the contours that are visible in its image. The second section deals with the image of a single generalized cone, and the third with objects that are composed of several cones joined together in various ways. These methods will be successful provided that no axis appears severely foreshortened in the image. Vatan (1976) and Vatan & Marr (1977) exhibit algorithms based on the theory, and the results of applying them to a number of natural images.

The main body of this article attempts to set out the motives and results of this research in plain English. I hope that it will be accessible to the general reader. The appendix contains precise statements of the restrictions and theorems, and gives their proofs. Most of the arguments there are geometrical, and a specialized mathematical background is not necessary to understand them.

Notation

As far as possible, I shall adhere to the following conventions. Surfaces, curves and lines in three-dimensional space will be denoted by upper-case Greek letters (Σ, Γ); curves and lines in an image from viewpoint V will be denoted by suffixed upper-case Roman letters (S_V, C_V); and upper-case Roman letters without a suffix (P, Q) denote points, either in 3-space or in the image. Lower-case letters obey the usual conventions, so that f, g, ι and ρ are functions, α, θ, ϕ and ψ are angles and x, y, z are coordinates.

1: The basic hypotheses and their implications

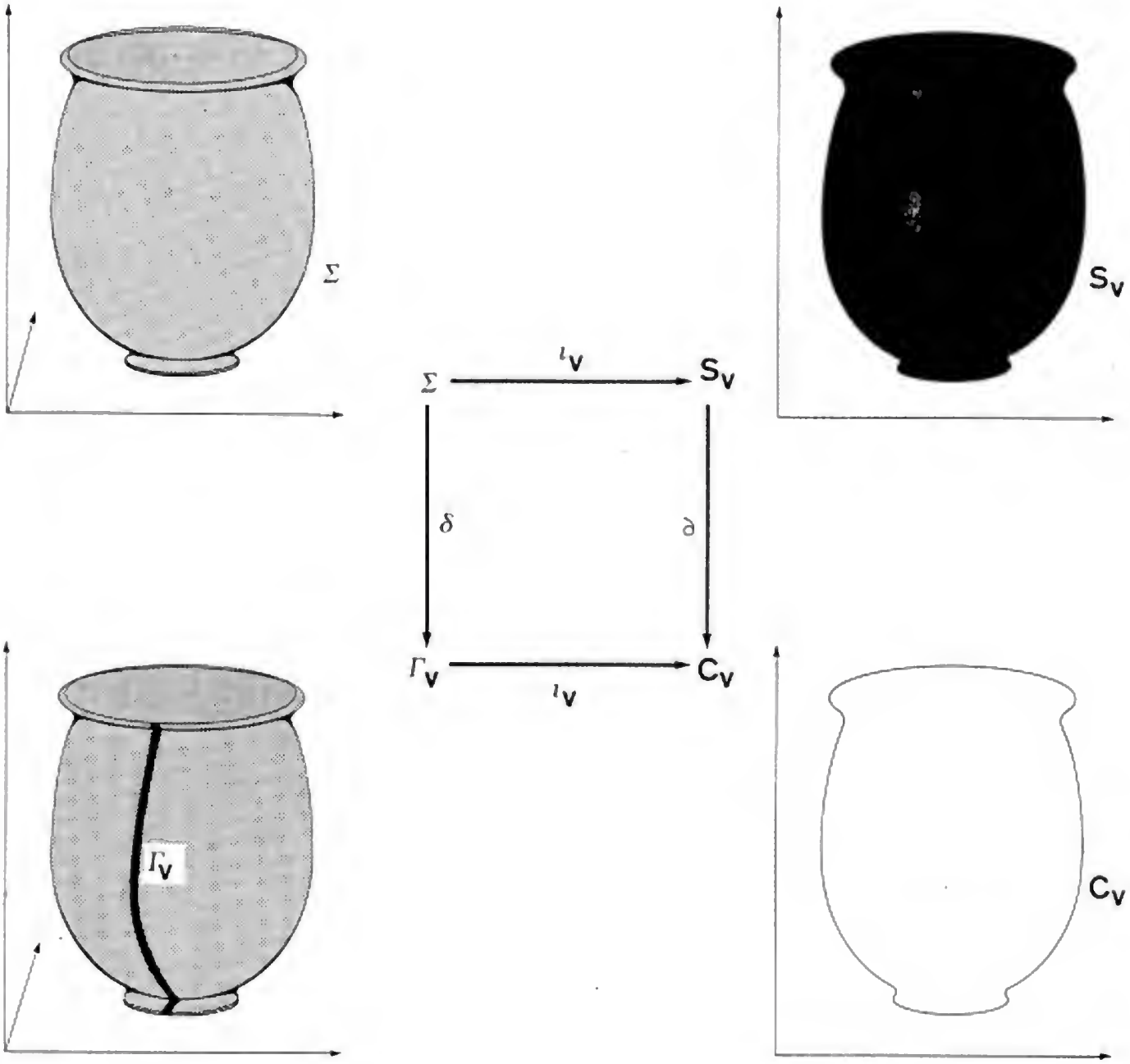
Our discussion will centre on the four structures that appear in figure 2. These are (1) some 3-D surface Σ; (2) its image or silhouette S_V as seen from a viewpoint V; (3) the bounding contour C_V of S_V; and (4) the set of points on the surface Σ, that project onto the contour C_V. We shall call this last set the *contour generator* of C_V, and we shall denote it by Γ_V.

We express the assumptions, to which I referred in the introduction, as restrictions on portions of the surface Σ or its image.

Restriction R1: The surface Σ is smooth.

We make this restriction only because we need to be able to distinguish convex pieces of contour from concave ones, and it is easiest to do this if the surfaces and contours in question can be differentiated twice. (In fact, we could have allowed the contours in the image to be composed of straight line segments, since the notions of convex and concave have a well-defined meaning in such cases; but little is gained by doing this).

Restriction R2: Each point on the contour generator Γ_V projects to a different point on



2. The three-dimensional surface Σ, viewed from a point V, forms the silhouette S_V in the image *via* the imaging process ι. The boundary of S_V, obtained by the boundary operator ∂ is denoted by C_V and we call it the contour of Σ. The set of points on Σ that ι maps onto C_V we call the contour generator of C_V, and it is denoted by Γ_V. The map from Σ to Γ_V induced by ∂ is denoted by δ.

the contour C_V.

This means that each line of sight from V to the edge of Σ -- that is, from V to Γ_V -- touches Σ at only one point, not at two points (as shown in figure 20) or along a line segment. The condition that each line of sight touches Σ at one, rather than at two or a finite number of points, is equivalent to saying that Σ is convex, as seen from this viewpoint (see figure 20). This is not as strong a condition as it appears at first sight, because in practise it will not usually be imposed for all viewpoints (e.g. theorem 1), and there are ways of regaining the generality that it excludes (theorem 5). Forbidding the line of sight from touching Σ along a line segment (as can happen for example if one views a cube from a direction parallel with one of its faces) is only a technical restriction; one can escape it without changing the situation in an important way by deforming the viewed surface very slightly.

Restriction R3: Nearby points on the contour C_V arise from nearby points on the contour generator Γ_V.

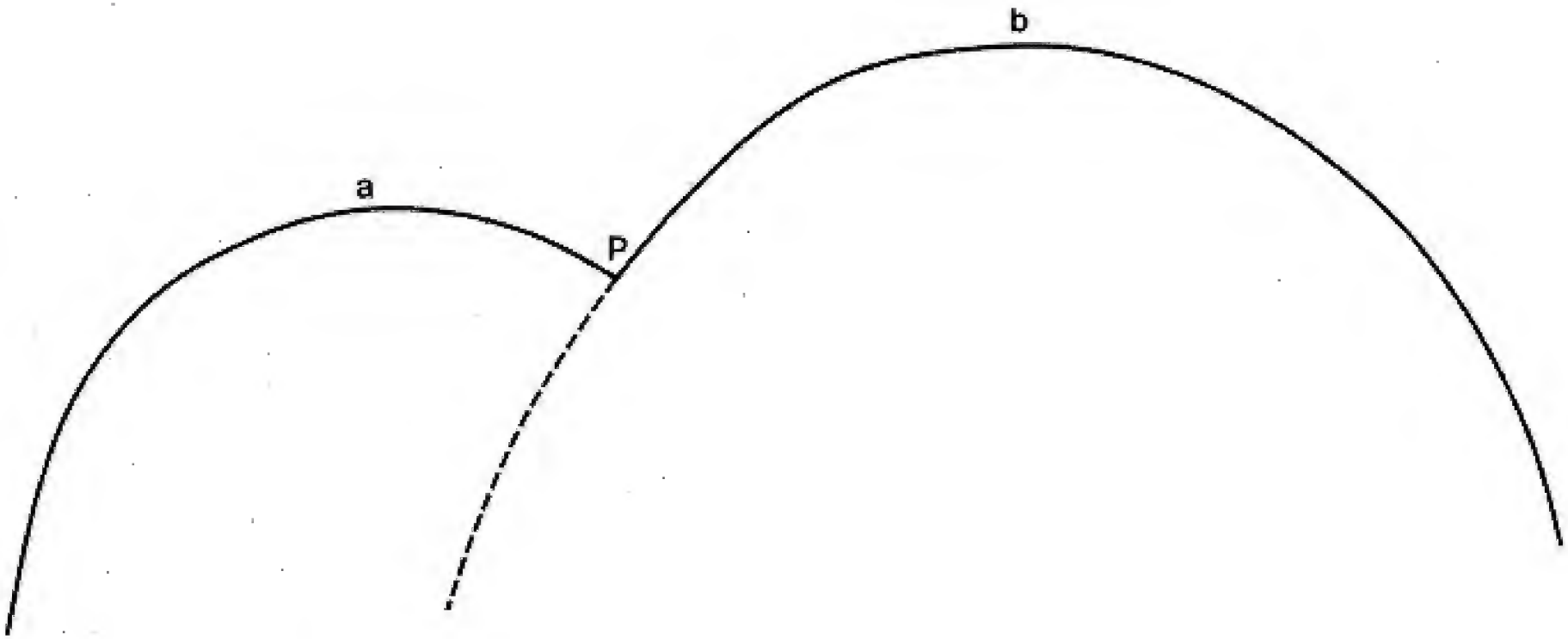
This condition is a powerful one, and is best explained by figure 3. Suppose that the contour *ab* of figure 3 really arose from two hills, but the dotted portion of *b* happened to be invisible. Then the contour generator of *ab* would be discontinuous at *P*, where it leaps from one hill to the next. This is the situation that *R3* forbids, and it is essentially equivalent to assuming that the image contains no invisible obscuring edges. Together, *R2* and *R3* imply that the contour generator Γ_V is a continuous curve across Σ – i.e. that it does not jump erratically from place to place on Σ. *R3* is a strong condition, but without it one can say almost nothing about Σ and I know of no way to proceed without it. Fortunately it is obeyed by most real-world images.

Remark

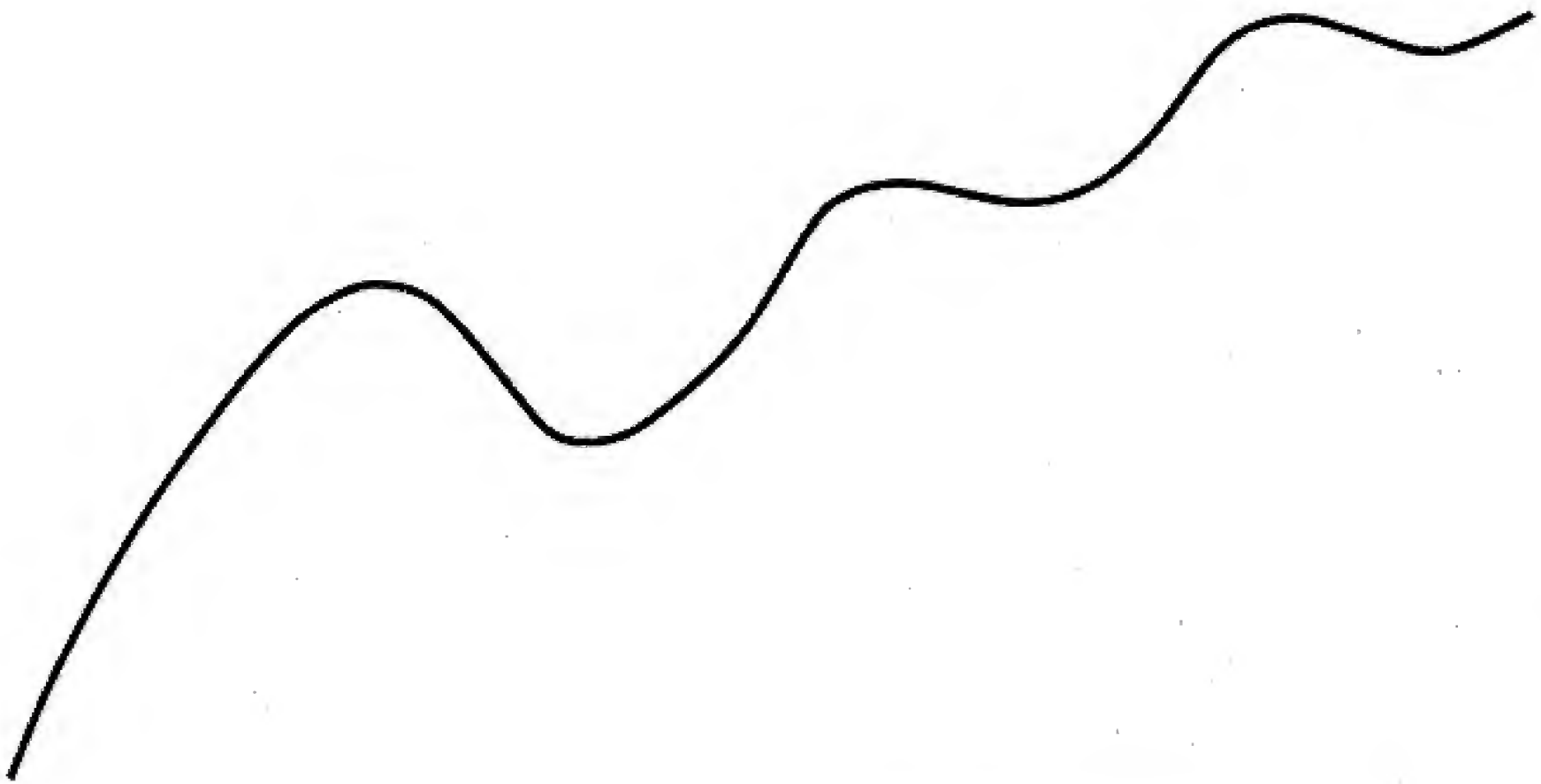
In fact, *R2* and *R3* are not quite independent, since if one assumes that the surface Σ is bounded, *R3* is a consequence of *R2*. To see this, notice that at points like *P* in figure 3 where *R3* is violated, the viewing ray to *P* grazes both hills, and so causes a violation of *R2*. Nevertheless, the two restrictions have sufficiently different meanings to make it worth stating them separately.

Using points of inflexion

The restrictions *R1-R3* are very general, and guarantee only the integrity of C_V and Γ_V, not their interpretability. Let us therefore suppose that a contour C_V, like that shown in figure 4, was obtained under conditions that satisfy these restrictions, and enquire what properties of C_V we can rely upon. Clearly, no metrical properties of C_V can be used, because C_V arises from viewing a surface Σ at an unknown orientation -- i.e. through at best a linear operator, and such operators do not preserve distances. The values of C_V's maxima and minima, and their separation, remain uninformative until substantially more is known about Σ and the perspective from which it is being viewed. But the qualitative



3. Restriction *R3*, which states that nearby points on a contour arise from nearby points on its contour generator, is essentially equivalent to assuming that there are no hidden obscuring edges. If *a* and *b* are the outlines of two hills, and if the dotted portion of *b* were invisible, *R3* would be violated at the point P where the contour generator leaps from one hill to the other.



4. A typical piece of contour. Until we have some information about the nature and position of the surface that gave rise to it, the only features that we can use are its points of inflexion.

notions of maxima and minima on a planar curve *are* preserved by a linear operator – that is, the distinction between convex and concave is invariant. This fact is captured by Lemma 1 of the appendix.

Let us therefore suppose that we have been presented with a contour segment like that shown in figure 4. Restriction *R3* guarantees that adjacent points on the contour arise from adjacent points on the surface Σ, but no metrical features are yet reliable. The only straightforward feature that remains is the distinction between a convex contour segment and a concave one, which rests in turn on the notion of an inflexion point. For a general surface Σ and contour generator Γ_V, even points of inflexion in C_V will often be meaningless, and to attribute significance to them is to make an additional assumption about Σ. So we next ask, how exactly should we formulate the assumption that points of inflexion are significant?

The restrictions *R1-R3* allow us to think of Γ_V, the contour generator of C_V on Σ, as a smooth piece of wire bent in 3-space. For inflexion points on C_V to be significant however, lemma 1 (see the appendix) tells us that we need two things; (1) the transformation due to the imaging process that produces C_V must be linear, and (2) the curve on which that transform acts must lie in a plane. Because the general perspective transformation is not linear, condition (1) tells us that our whole theory applies only to distant viewing points, because only in these conditions is the imaging process a linear projection. Condition (2) informs us that the convex-concave distinction can be meaningful in C_V only if the bent wire that is Γ_V lies in a plane. This gives us our fourth condition.

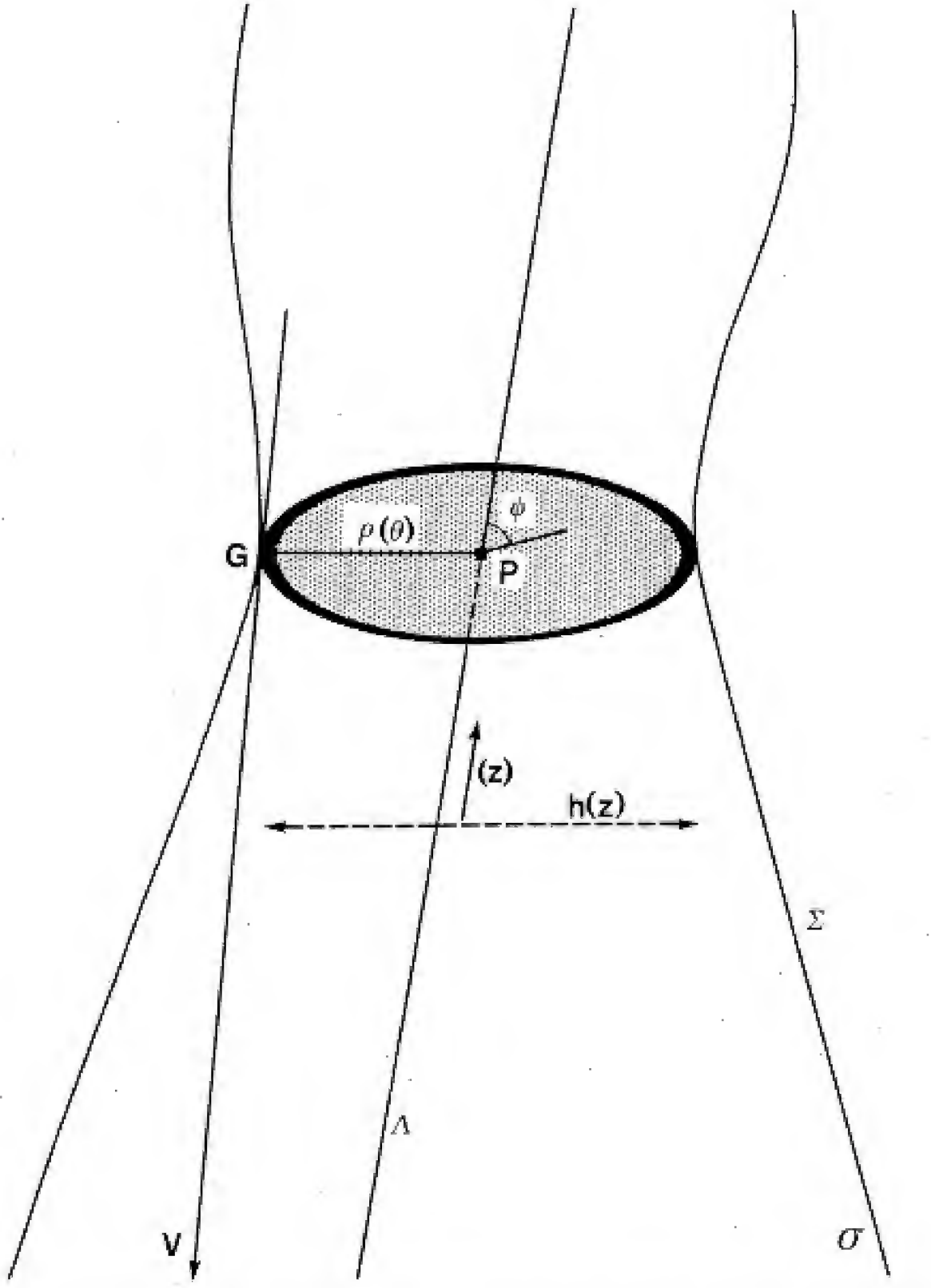
Restriction R4: The contour generator Γ_V of C_V is planar.

This condition is a strong one, and sharply delimits the class of admissible surfaces Σ. There seems however to be no way of avoiding it if one wishes to use the distinction between convex and concave contour segments.

Implications of the four restrictions

A *generalized cone*, illustrated in figure 5, is defined to be the surface swept out by moving a simple smooth cross-section along some axis, at the same time magnifying or contracting it in a smoothly varying way. This cross-section is defined by the function $\rho(r, \theta) = 0$, and when the cross-section is convex, we shall use cylindrical coordinates $r = \rho(\theta)$. The magnification of the cross-section at each point is specified by the function $h(z)$, where z is the distance measured along the cone's axis. The axis itself will be labelled Λ. Notice that in general the z axis need not be perpendicular to the plane $z = 0$ of the cross-section. These conventions are illustrated by figure 5.

We may demand that the restrictions *R2* - *R4* hold for all views, or for a subclass of the possible views of Σ. If we demand that they hold for only one (distant) viewpoint, this imposes no interesting restrictions on the nature of Σ. Theorem 1 studies the two dimensional case, when the restrictions are assumed to hold for all distant viewpoints whose



5. The definition of a generalized cone. In this article, a generalized cone is the surface generated by moving a smooth cross-section ρ along a straight axis Λ. The cross-section may vary smoothly in size (as prescribed by the axial scaling function $h(z)$), but its shape remains constant. The eccentricity of the cone is the angle ψ between its axis and a plane containing a cross-section.

lines of sight lie parallel to some fixed plane, and it is the most interesting result of the section. Finally, theorem 2 studies the consequences of assuming that the restrictions hold for *all* distant viewpoints.

The basic theorems

Earlier, we defined S_V to be the image of Σ as seen from the vantage point V (figure 2). This is equivalent to saying that S_V is the perspective projection of Σ from the point V. In theorem 1, we shall make two simplifying assumptions about the projection S_V; first that the projection is orthogonal, which is approximated when the vantage point V is very distant from Σ compared to its size, and second, that the viewing directions are confined to a plane Π round Σ and which intersects Σ. We deal in some sense only with "side" views of Σ, and are forbidding "end-on" views. Such projections are completely specified by the direction of the vantage point from Σ in the confining plane Π, and we denote this by the angle ϕ. We shall use the notation S_ϕ, C_ϕ and Γ_ϕ in place of S_V, C_V and Γ_V to indicate that the above restrictions are in effect. The proofs of theorems 1 and 2 are set out in the appendix. I give here their statements in plain English.

> *Theorem 1.* Σ is a generalized cone with convex cross-section if and only if *R1* is satisfied, and *R2* - *R4* are satisfied for all orthogonal projections ϕ associated with some plane Π, in the sense defined above. This plane lies parallel to the cross-section of the cone.
>
> *Theorem 2.* Σ satisfies *R1* and *R2* - *R4* for *all* distant vantage points V if and only if Σ is a quadratic surface.

Remarks about theorems 1 and 2

It is theorem 1 that allows the crucial step for the overall argument. It says that if, for distant viewpoints whose viewing directions lie parallel to some plane, a surface's shape can successfully be inferred using only the convexities and concavities of its bounding contours in an image, then that surface is a generalized cone with convex cross-section, or is composed of them. Hence if one *assumes* that one can discover a surface's shape from such information, then this is equivalent to assuming that the viewed surface is a generalized cone. The assumption of theorem 1, about orthogonal projections parallel to the plane of the cross-section, is tolerable because as we shall see the methods to which the theory gives rise usually degrade only slowly as one moves nearer, increasing the effects of perspective, or out of the plane of the cross-section. Furthermore there does appear to be something special about the perception of views that look down the z-axis of the figure (see the remarks made by Marr & Nishihara (1977) about Warrington & Taylor's (1973) "unconventional views").

Theorem 2 is interesting, because it shows how very strong our restrictions are. One can gain a feel for how the planar condition *R4* fails for higher-order surfaces by studying the behaviour of $x^{2n} + y^{2n} + z^{2n} = 1$ (see figure 6). This surface is a sphere for $n = 1$, and tends to a cube as n grows large. The contour generator Γ_V, which is a circle for n

- I (figure 6a), becomes the outline marked with thick lines in figure 6b for high values of n. This contour-generator is clearly not planar; as n increases, the lower third of the contour generator is pulled towards the viewer, and the upper third is pushed away.

These results provide a further argument for using something based on generalized cylinders (Binford 1971) for the internal representation of shape (see Marr & Nishihara 1977), an argument based not on utility, as most other justifications are, but on the assumptions implicit in the decoding of an image. It is indeed extremely fortunate that many important three-dimensional structures can be closely approximated by a few generalized cones, although it is not accidental that objects whose shape was achieved by growth like limbs and stalagmites, can be so approximated.

2: Interpreting the image of a single generalized cone

Theorem 1 essentially tells us that, when trying to infer the shape of a surface from its bounding contours in an image we cannot avoid assuming that the surface is a generalized cone. We are now faced with an obvious question. If we assume that our data consists of contours in the image of a generalized cone, how may we interpret them? To specify a generalized cone, we have to specify its axis A, cross-section $\rho(\theta)$, and axial scaling function $k(z)$; how can we discover them from an image?

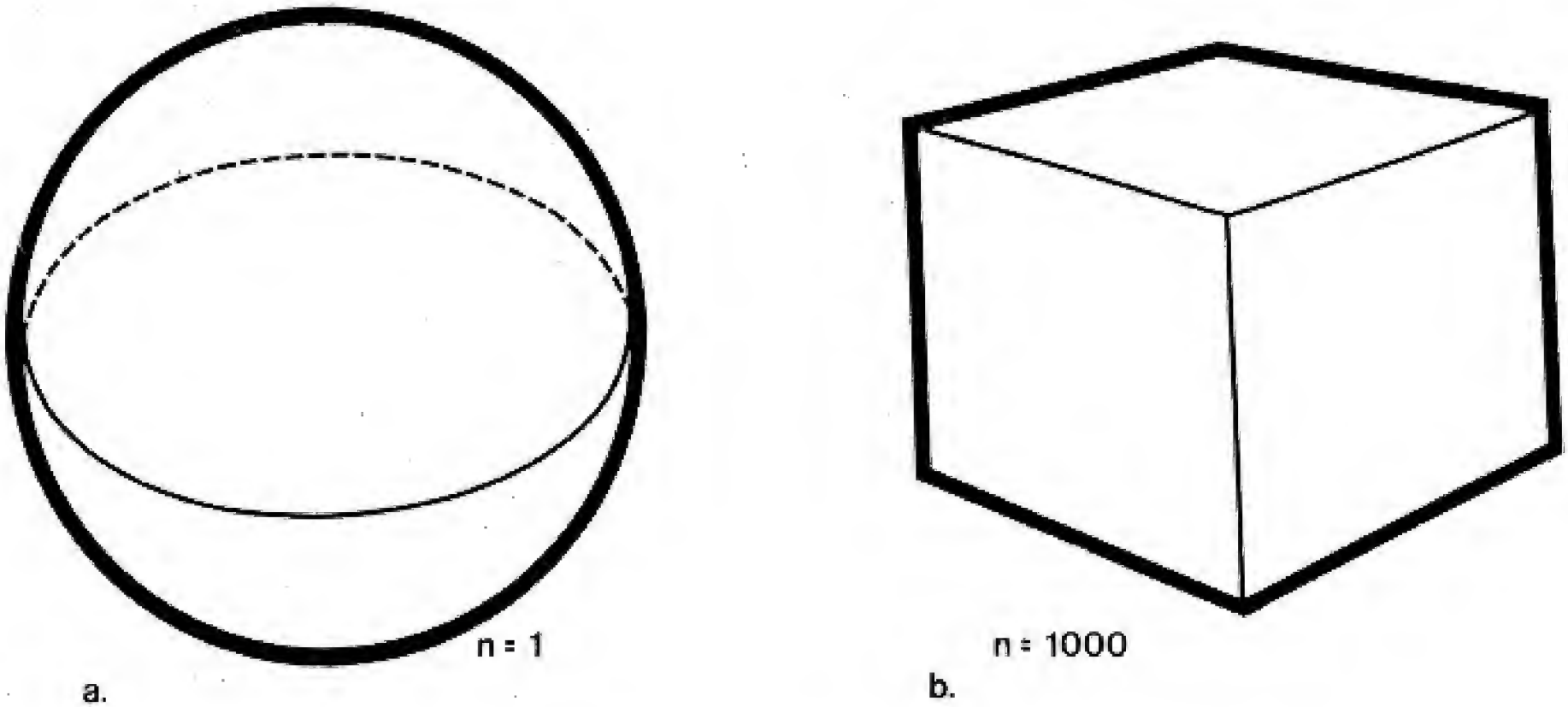
The answer to this question commences with theorem 3, which shows how the occluding contours in an image may be used to find the "image" of the cone's axis for those distant viewpoints that lie in the privileged viewing plane referred to in theorem 1. In general, of course, our viewpoint will not lie in this plane, and so we have to examine the stability of this result as the viewpoint moves out of the plane. This is achieved by theorem 4, which introduces a new concept called the *skeleton* of a generalized cone. The skeleton is not a difficult idea, however, since it is very like the set of lines a cartoonist draws to convey the shape of a curved object. The idea of a skeleton allows us to extend the theory to generalized cones whose cross-section is not convex. Requiring the restriction *R2* to hold for all ϕ-projections essentially forbids this class of cones, and I said earlier that one can circumvent this restriction in practise. Theorem 5 shows how. Finally, there is a short discussion about cases in which the cone is viewed from a nearby rather than from a distant point, and cases in which the axis of the cone is not a straight line.

The overall purpose of the section is to give a set of methods for interpreting the image of a single generalized cone. The methods derived here will not succeed for all views; they will fail when the image of the cone's axis is substantially foreshortened. It is part of the overall theory that such views have to be handled differently (Marr & Nishihara 1977).

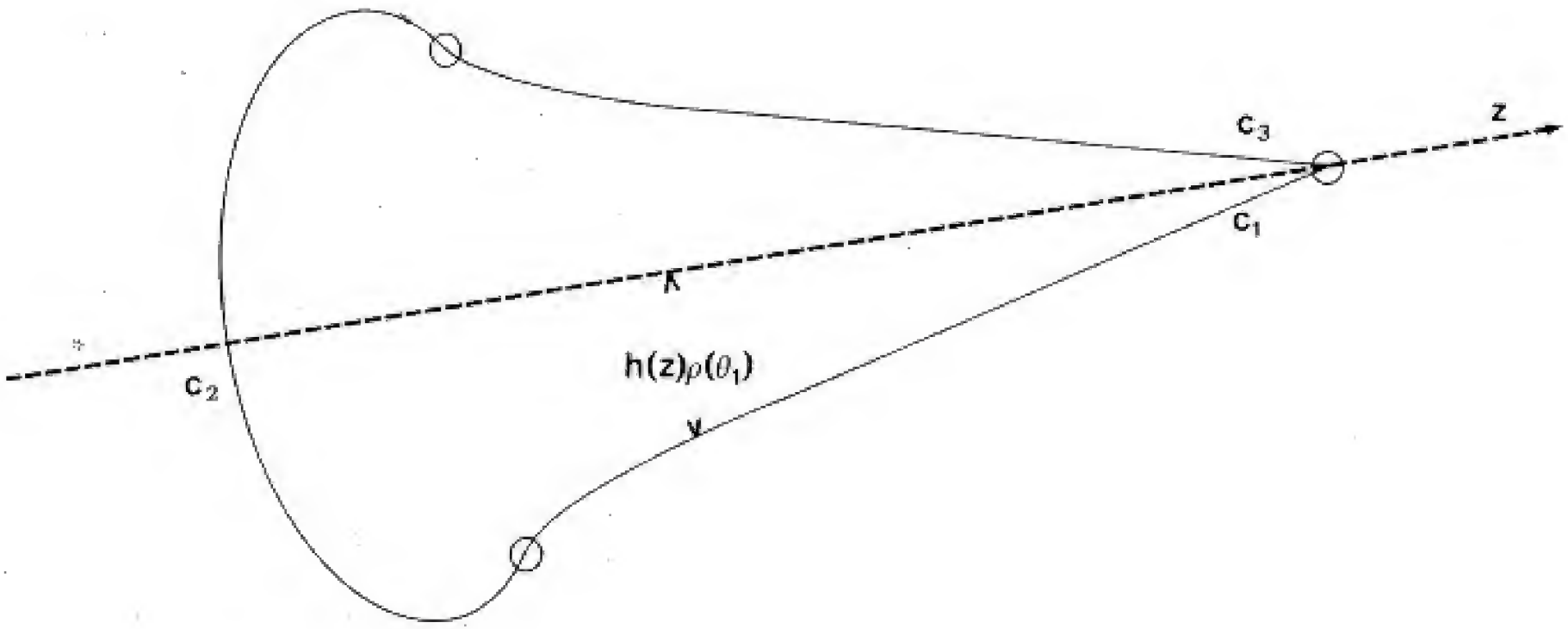
Finding the axis from a favourable view

Provided that the viewed surface is a generalized cone, and that the viewing point satisfies the conditions of theorem 1, the axis of the cone may easily be determined by the rough symmetry formed around it.

Theorem 3 (Axial Symmetry). Let Σ be a generalized cone with convex cross-



6. The surface $x^{2n} + y^{2n} + z^{2n} = 1$ for $n = 1$ (figure 6a) and $n = 1000$ (figure 6b). The contour generator is shown in thick lines. It is planar (a circle) in 6a, but not for $n > 1$. Figure 6b shows how the contour generator is pulled out of a plane for high values of n.



7. The image of a generalized cone whose axial scaling function k contains at least one concave segment. Theorem 3 establishes the existence of a qualitative symmetry between at least the concave components of the contour, which here are labelled C_1 and C_3. The axis of the symmetry A^6 is shown dotted in the figure, and it is the image of the axis of the generalized cone.

section $\rho(\theta)$, and let the cross-section scaling function $h(z)$ contain at least one concavity. Then for all viewing directions ϕ

(i) the silhouette of Σ, C_ϕ, decomposes into $n > 2$ contour segments by splitting it at points of inflexion;

(ii) the image of the axis Λ of Σ establishes an axial symmetry between at least $(n - 2)$ contour segments, including all concave ones. Corresponding segments are either both convex or both concave;

(iii) the ratios of the distances of corresponding segments either side of the axis of symmetry are all the same.

Corollary: The image of the axis of Σ is uniquely determined if there exists only one such axial symmetry.

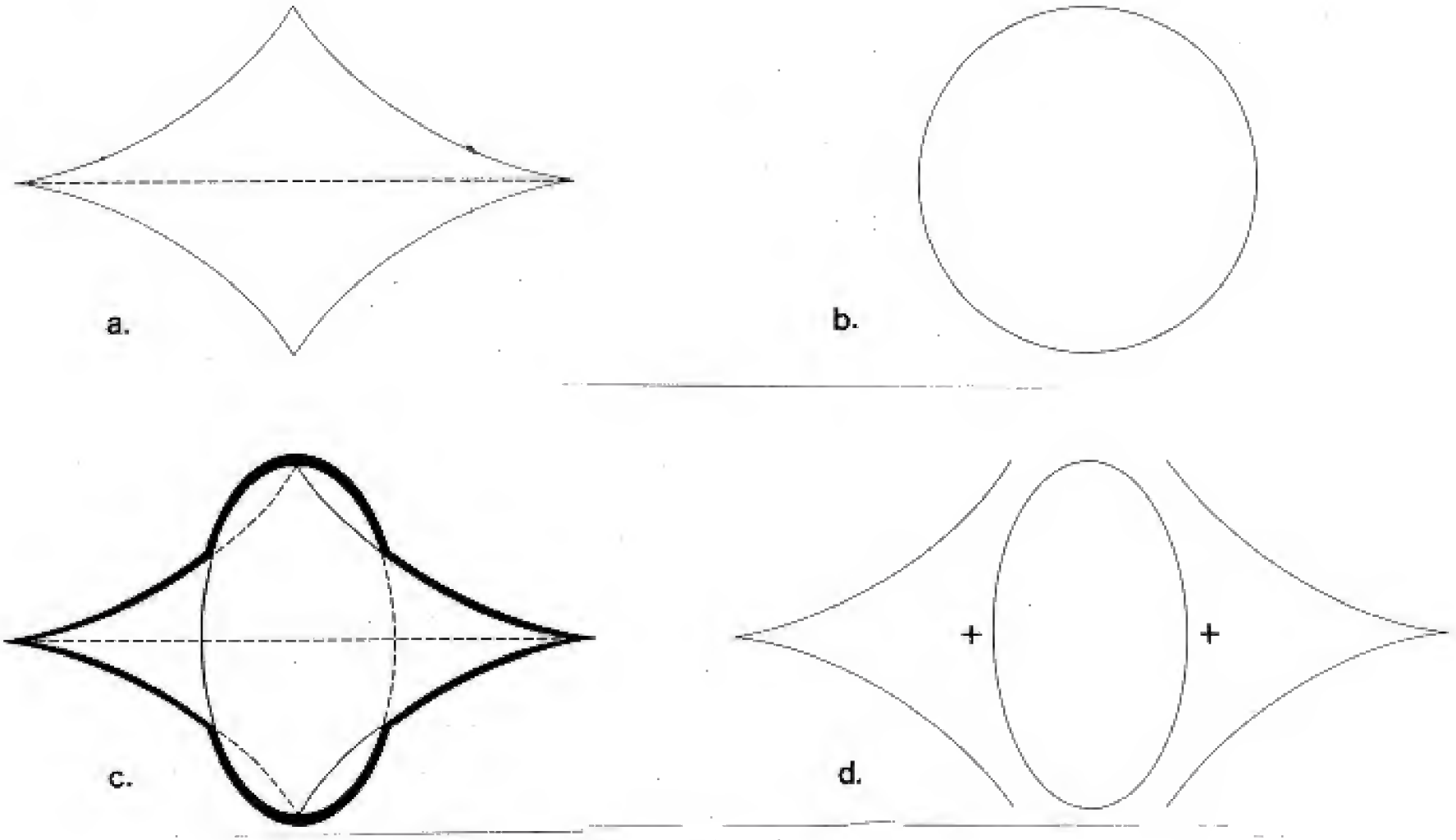
This theorem is best explained by looking at figure 7. Here, we see that the contour divides at inflexion points into three segments, labelled C_1 to C_3. The two concave segments C_1 and C_3 are roughly symmetric about the image of the cone's axis Λ, although their distances away from the axis may not be equal. The third clause of the theorem states that, if C_1 is half as far from Λ as C_3, then the same will be true of all other segments that correspond under the symmetry. We shall call the type of symmetry established by theorem 3 a *qualitative symmetry*. Its important features are (a) that it holds between convex or concave *segments* of a contour, and (b) that it includes a scaling factor.

Although there is a point-wise symmetry between the contour on the two sides of the axis, unless the viewed surface is a right generalized cone and the contours are faithfully diagnosed in the image, such symmetries are expensive to detect. A qualitative symmetry, on the other hand, does not have to be found on a point-by-point basis. This is important because it makes *finding* the symmetry, and hence its axis, a practical computational proposition. By dividing the contour into convex and concave segments and noticing that the symmetry preserves this distinction, we have greatly reduced the number of items that have to be examined and made the computational load acceptable.

There is one other point of importance about this result and it comes from the corollary, which says that if only one symmetry exists among the contours, the axis of Σ is determined uniquely. This means that the analysis of contour is self-checking, and one does not have to appeal to the "familiarity" of the deduced shape to know that one has a valid interpretation of the image. This is of course essential if one is to be able to analyze novel shapes. The reader will observe that all of the theorems, that are directed at the analysis of contour, have uniqueness corollaries like that of theorem 3. It is on these that the algorithms themselves will rest most directly.

Viewing directions not coplanar with a cone's cross-section

We next ask what happens to the generalized cones of theorem 3 if the viewpoint remains distant, but if the viewing direction moves out of its constraining plane Π? As long as the image S_γ approximates an orthogonal projection parallel to the plane of the generalized cone's cross-section, variations in the silhouette of Σ are due to changes in the scaling function $h(z)$ along the cone's axis Λ, as illustrated by figure 8a. On the other hand,



8. While the viewing direction remains distant and coplanar with a generalized cone's cross-section, variations in the contours in the image are caused solely by variations in A, the scaling function (figure 8a). When the viewing direction is distant and lies parallel to the axis of the cone, the contour in the image is caused solely by ρ, the cone's cross-section (figure 8b). In between, the contour is approximately caused by a mixture of the two (figure 8c). The bounding contour shown thickened in 8c violates restrictions *R3* and *R4*, but it can be decomposed into components that do satisfy them (figure 8d), to which the machinery of the theory can then be applied.

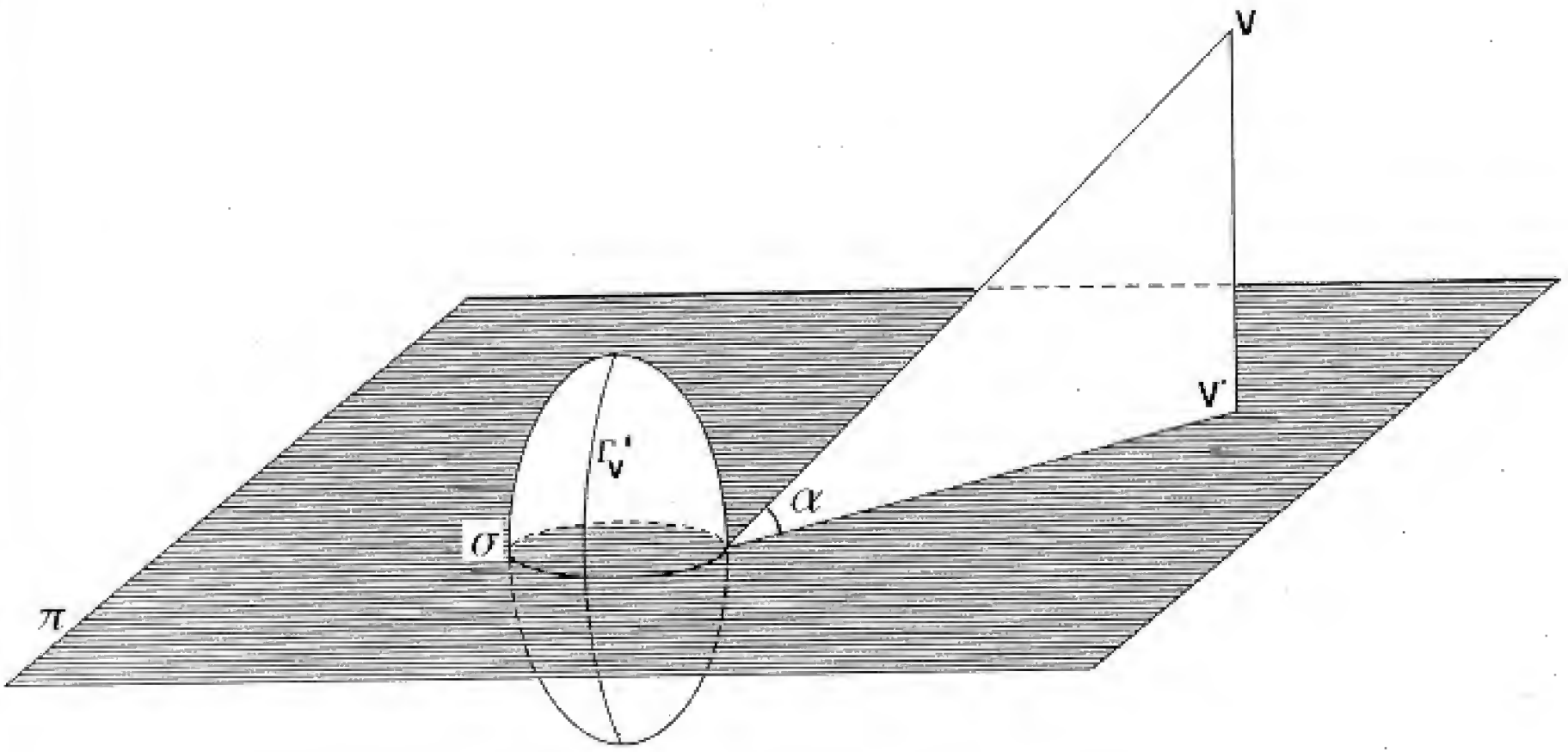
when the viewpoint is moved so drastically that the viewing direction lies parallel to $\mathbf{A}$, the silhouette of Σ is due entirely to $\rho(\theta)$, its cross-section function (figure 8b), and is in fact due to the cross-section of Σ at the point where $h(z)$ achieves its maximum value.

At both of these extremes the contour generator Γ_V of Σ is planar, but for other viewing directions it need not be. This is obviously true for a dumbell shape, or for an ice-cream cone, but theorem 2 assures is that it is more insidiously true even for a surface as simple as an egg, where the contour generator is a circle for the end-view, a near oval for a side view, and slides from one to the other in between (see figure 10b below). Another example is the surface shown in figure 8c; the contour generator from this view is clearly not planar.

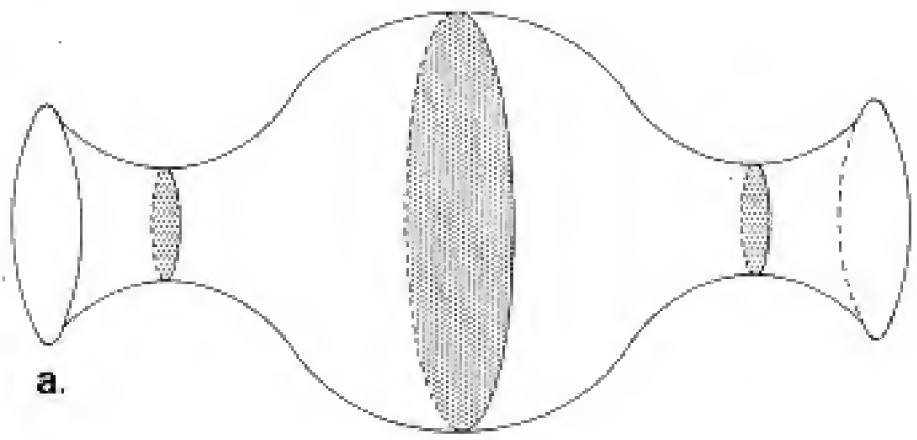
How are we to handle such views? For convex objects like an egg, where the only occluding contours arise from its silhouette, there is very little more one can do to infer its shape when seen from one of the intermediate views, unless one knows something about the orientation of the egg relative to the viewer. For objects that are not convex, like the double spike of figure 8, one can separate the contours that arise in the image into two classes; those approximately due to the "sides" of the figure (the two spikes separated in figure 8d), and those approximately due to a cross-section, like the central ellipse in figure 8d.

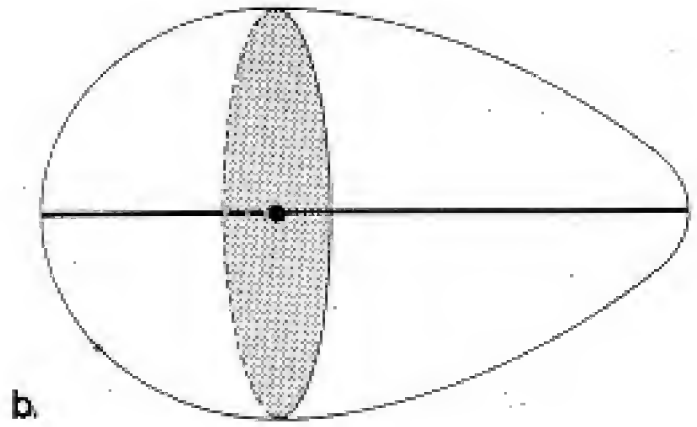
This division gives us our main tool for analyzing non-standard views, and it is best explained with the help of figure 9. Suppose that a generalized cone Σ is being viewed from a distant point V and the line of sight is not parallel to the plane of the cone's cross-section. The contour generator for viewpoint V is approximated by two components. One is easy to define; it is the places on Σ where the size of the cross-section is stationary — that is, where $h(z)$ achieves a maximum or minimum. For an egg, it is the fattest cross-section, and other examples are shown shaded in figure 10. We call these curves *radial extremities*, and denote them by $\mathbf{T}$ — notice there is no suffix, since $\mathbf{T}$ does not depend on the vantage point. The other idea we have to make as precise as possible is what we mean by the "sides" of Σ from a viewpoint such as V, and for this we make the construction illustrated in figure 9. We drop a perpendicular from V to V', which does lie in the standard viewing plane. Then the contour generator for viewpoint V is approximately the projection of the contour generator $\Gamma_{V'}$ for V', which is simply Γ_ϕ for some angle ϕ. For example, in the case of an egg of length l and diameter d, the skeleton (shown in figure 10b) has length $l.sin(\chi)$ and width d, when viewed in a plane containing the egg's major axis at an angle χ to it. For angles where $l.sin(\chi) > d$, this is a reasonable approximation, and when $l.sin(\chi) < d$, we have an "unconventional" view.

The reason why the skeleton is a useful construct for recognition is that one can detect its presence in an image by the many relationships that exist among its parts. In fact, we can use these relationships to set up constraints on a set of occluding contours such that if those constraints are all satisfied by a unique interpretation of the contours in the image, we can be reasonably certain that we have found a skeleton, and hence can interpret the contours as arising from a generalized cone Σ whose axis is then determined. The relations themselves consist of qualitative symmetries and parallelism, and are preserved by an orthogonal projection. Hence provided that the contours as seen from vantage point V in figure 9 are approximately the projection of the contours as seen from V', the relations will



9. Contours that arise in views of Σ from directions that do not lie in the plane of Σ's cross-section are approximated by two components. One component consists of local maximum and minimum cross-sections of Σ, called its radial extremities Υ, and the other component is the contour generator for the nearest viewpoint whose viewing directions do lie in the plane of the cross-section of Σ. For a viewpoint V, this specially constructed viewpoint is called V', and it is illustrated in the figure. The shaded plane indicates a plane of Σ's cross-section. The set of contour generators obtained by this approximation is called the skeleton of Σ for viewpoint V.





10. Examples of skeletons of surfaces, as defined in the text and in figures 8 and 9. The cross-sections responsible for the radial extremities of the surfaces are shown shaded. The skeleton ceases to be a reasonable approximation to the contours that occur in the image whenever the viewing angle is such as to make the projection of the length of the cone less than the orthogonal projection of its width. For such views, the methods of this article will fail.

still hold in the image formed from V.

> *Theorem 4 (Skeleton Theorem).* Let $\Gamma_{V\prime} \cup \mathbf{T}$ be the skeleton of Σ associated with some vantage point V. Then provided that C_V can be thought of as being formed by the orthogonal projection of $\Gamma_{V\prime}$ along the direction to the vantage point V,
> (i) C_V is qualitatively symmetric about the image of the axis Λ of Σ, in the sense of theorem 3
> (ii) the image of $\mathbf{T}$ consists of one or more connected components, through which Λ passes, and between any two of which there exists a mapping that is *(1 - 1)*, continuous and onto, that preserves the gradient of the image of $\mathbf{T}$ at each point.

Corollary: If Σ obeys the prerequisites of theorem 4, and if the image of its skeleton decomposes in a unique way into two components that satisfy conditions (i) and (ii) of the theorem, then these components are the images of Γ_V and of $\mathbf{T}$. The axis of symmetry of the image of Γ_V is the image of the axis of Σ.

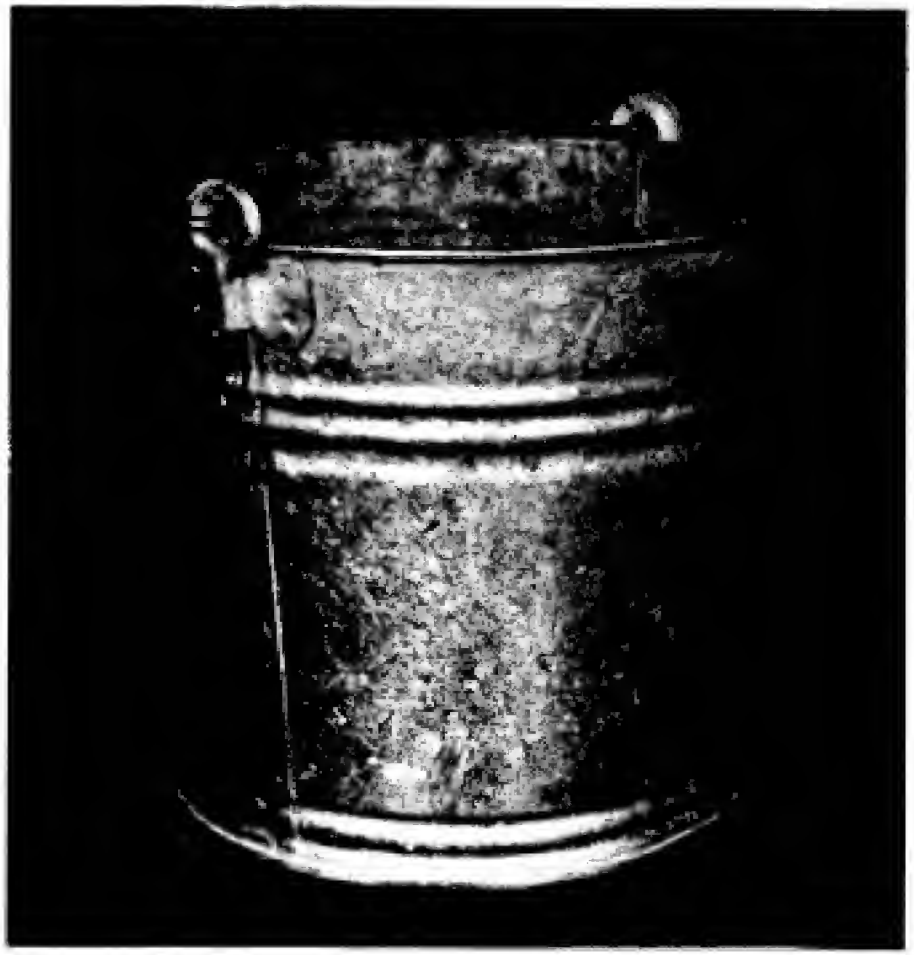
This theorem makes explicit the many relations between the elements of a skeleton's image, and its practical importance is illustrated by figure 11. The theorem states that the image of the "sides" obeys quite well the symmetry relation of theorem 3, and one can see from the figure that this is true of the sides of the bucket in the image. The axis of symmetry of the sides is the axis of the bucket. The theorem also says that the images of a cone's radial extremities are all parallel to one another, and embrace the cone's axis. In figure 11, there is a clear parallel relationship between the image of the bucket's top, the corrugations in its side, and the visible part of its base.

As in the case of theorem 3, the diagnostic power of this result lies in the corollary. It does not guarantee that a given set of occluding contours can be interpreted, but if a unique interpretation exists that satisfies these conditions, then it will be correct. In a real image, many parts of a cone's skeleton will be obscured, but this hampers the finding of relationships like parallelism and qualitative symmetry only slightly. One can devise a cooperative algorithm (Marr & Poggio 1976) that can operate on the description of a contour to find relationships of this kind between its pieces (Vatan & Marr 1977).

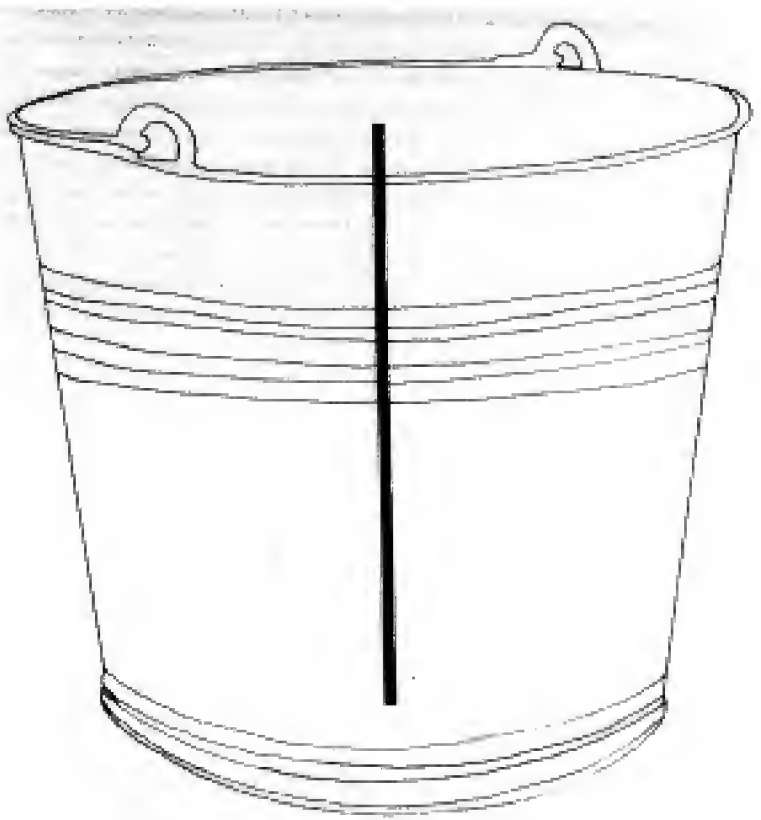
Generalized cones whose cross-section is not convex

We are now ready to extend the theory to the case where the cone's cross-section contains concavities. The important difference between this and the case where the cross-section is convex is that occluding contours can now also arise from local maxima and minima in the cross-section ρ. For example, in the image of a fluted pillar, there are many lines running parallel to the axis of the pillar, corresponding to the local maxima or minima in the pillar's cross-section.

This gives us the extra tool we need to extend the analysis of theorem 4. Contours that are due to convexities and concavities in the cross-section ρ behave like the fluting on a pillar, so we call them the cone's *fluting* and denote them by the letter Φ. The



a.



b.

11. Methods based on the corollary to theorem 4 suffice to solve this image of a bucket. An axial symmetry is established by its sides about the bucket's axis (shown thickened), and a parallel relationship holds between components of its radial extremity. Here, these are the bucket's top and bottom, and the corrugations in its side.

fluting on a cone with variable cross-section behaves rather like the silhouette of theorem 3. Convexities and concavities in the fluting on one side of the cone's axis are in qualitative symmetry with the fluting on the other side (as seen by the viewer). This means that contours in the fluting obey a set of qualitative symmetry and rough parallel relations among themselves, similar but orthogonal to those obeyed by the radial extremities. These relationships can be used to interpret the contours in an image, in a way analogous to theorem 4. The extension of theorem 4 to the case of a cone with fluting is theorem 5.

> *Theorem 5.* Let $\Gamma_V \cup \Upsilon \cup \Phi$ be the skeleton and fluting of Σ associated with some distant vantage point V. Then
> (i) C_V and Υ obey theorem 4.
> (ii) The image of each portion $\{(h(z).\rho(\theta_i), \theta_i, z)$, for fixed θ_i and varying $z\}$ of the fluting is either a straight line, or it divides into convex and concave segments that are in *(1 - 1)* correspondence with the convexities and concavities in that part of the image of Γ_V which lies on the same side of its axis of symmetry.

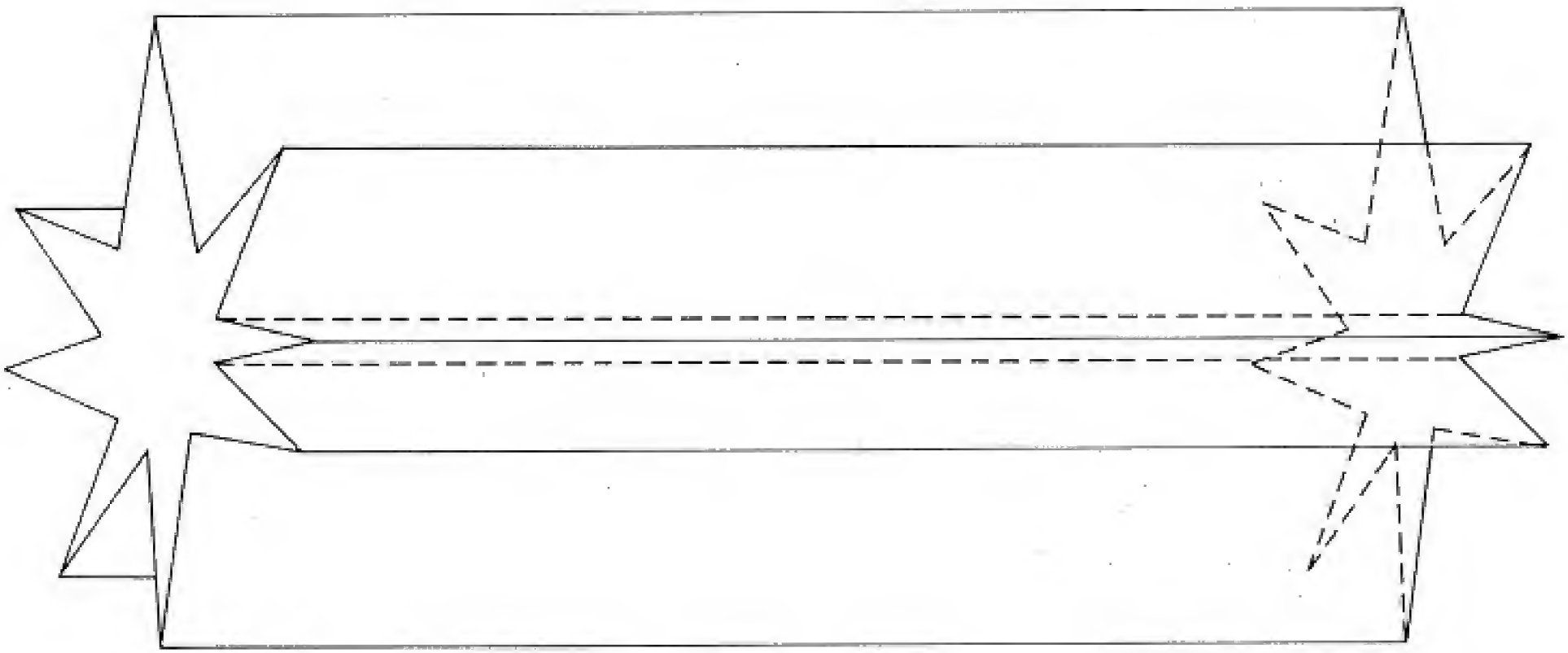
Corollary: If Σ is a generalized cone, and if the contours in its image decompose in a unique way into three parts that satisfy the conditions of theorem 5, then those parts are (i) the image of Γ_V, whose axis of symmetry is the image of the axis of Σ; (ii) the image of Υ, the radial extremities of Σ; and (iii) the image of Φ, the cone's fluting.

Once again this result enables us to set up a system of constraints on the contour present in an image such that, if the constraints are satisfied by a particular labelling of the contour, that labelling enables us to discover the axis A of Σ, and other information about its cross-section ρ and axial scaling function h. The algorithms that implement this method need only recognise the properties of parallelism and qualitative symmetry between a small number of elements. This result reaches slightly beyond the scope of this article since it deals with contour that is not necessarily occluding. It also extends naturally to the case where ρ contains creases (points of discontinuity in gradient), which is helpful because creases often give rise to edges and highlights in an image.

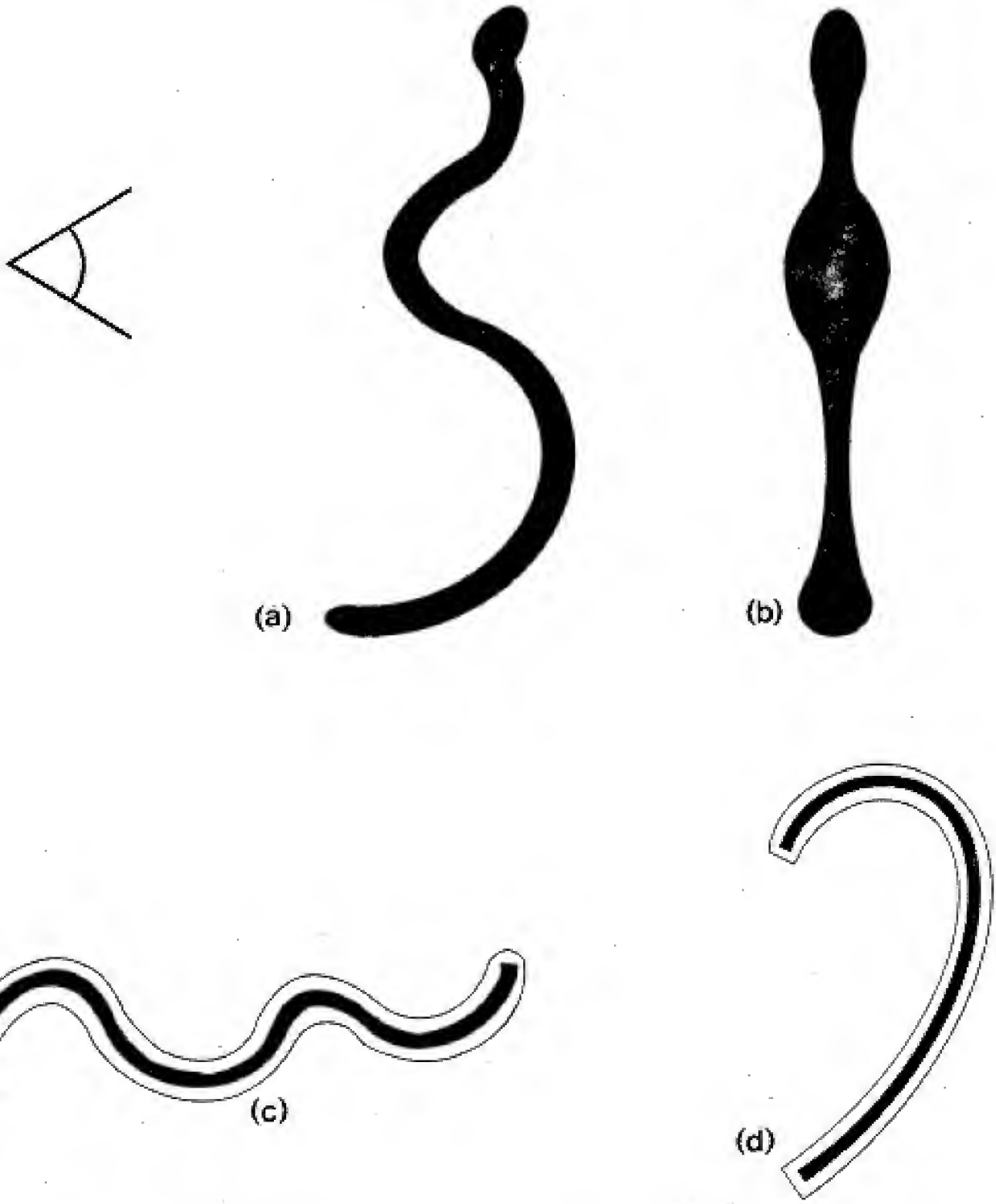
Nearby viewing points and curved axes

The methods discussed in this article are ill-suited to images that arise from nearby viewing points, and are of little use for cones with curved axes unless their cross-sections are simple. These points are best made by figure 13, which shows a serpent weaving towards and away from a nearby viewer (figure 13a), who sees an image that resembles figure 13b. The points of inflexion in figure 13b are caused by perspective, and to recognise this one needs other cues, like texture gradients and stereopsis.

If one sees the contour that appears in figure 13c, one can and does infer the shape of a snake. Cases such as figures 13c & d, where the scaling function h is roughly constant, are easy to deal with; so are other cases where the qualitative symmetry of theorem 3 is reversed (i.e. convex segments match concave segments, not convex ones), but in general the situation can be complex. I have been unable to derive any substantial results from



12. When the cross-section of Σ is not convex, we need to add the images of its maxima and minima in order to arrive at a complete model of its visible contours. The lines traced out by these maxima and minima are called the surface's fluting, and they are added to the skeleton as defined in figures 9 and 10 to produce what is called the surface's complete skeleton. This figure depicts a surface with fluting.



13. The methods of this article are based on the distinction between convex and concave contour segments. They are therefore unsuitable for images of nearby objects. For example, if a viewer is close to a snake (as in 13a), the image he sees will be something like 13b. The convexities and concavities in this are mostly due to the perspective transformation, and they do not reflect properties of the viewed surface. The figures in 13c and d are generalized cones with curved axes. It is not known how to deal with these except in simple cases like those depicted here.

circumstances in which the surface Σ and its viewing point are unconstrained.

3: Surfaces composed of two or more generalized cones

We have hitherto been concerned with the appearance of a single generalized cone. Real-life objects are often approximately composed of several different cones, joined together in various ways (see Marr & Nishihara 1977 figure 6), and we therefore have to study ways of decomposing a multiple cone into its components -- for example, a human body into arms, legs, torso and head. The way in which two cones join has a profound but usually local effect on the contours produced by the resulting surface, and it can upset the qualitative symmetry and parallelism on which our earlier results depended by interfering with the inflexion points on which primary contour description is based. Therefore, the algorithms for interpreting occluding contours in an image must incorporate a sensitivity to situations that can arise as a result of joins.

In this section, we study the common types of cone-cone junction, classify the appearances to which they can give rise, and indicate how algorithms for their detection may be constructed. In order to do this, we once again have to place some restriction on the way in which a join is configured. The one that I choose is:

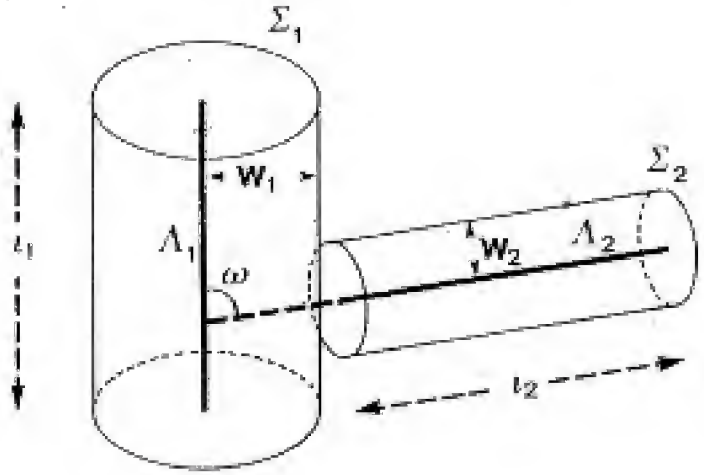
Restriction 5: The axes of two joined generalized cones are coplanar,
which enables one to relate the silhouette of the junction between two cones to the angle between their axes and their axial scaling functions. If the two axes are not coplanar, the surfaces at the junction are rather unconstrained. In practise, *R5* is not a severe restriction. Provided that the two axes approach one another closely relative to the width of their respective cones, the coplanar condition will be satisfied closely enough.

A: Side-to-end joins between two generalized cones

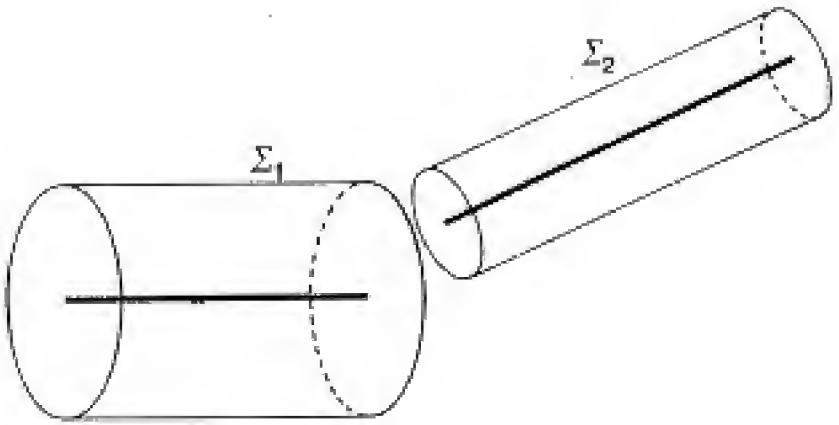
The most useful common feature of the join between two cones is that it gives rise to one or two deep concavities in the surface's silhouette. This feature is unfortunately not a necessary concomitant of a cone-cone junction, and although it plays a large part in our algorithms for detecting such a junction (Vatan 1976), its role in the underlying theory is surprisingly slight.

It is convenient to divide the types of join that can occur into two classes, those in which the end of one cone is attached to the side of the other, and those in which the two cones are attached at their ends. The two types of join are illustrated in figure 14, and a formal statement of the distinction between them is given in the appendix. These two cases are not quite exhaustive, but the intermediate cases introduce no new points of interest.

From the point of view of diagnosing these joins, the important difference between them is that there are often two concavities associated with a side-end join, (one on each side of A_2 as shown in figure 14a), but there need not be for end-to-end joins (figure 14b). We analyze the possible configurations case by case.



a.



b.

14. The two main types of joins considered in this article. 14a shows a side-to-end join, and 14b shows an end-to-end join. In the case of a side-to-end join between two convex cones (like 14a), theorem 6 guarantees the presence of a minimal concave angle in the bounding contour, due to the join itself. Provided that the cones are long relative to their width, the total concavity will be near 180°.

A1: Both cones are convex

An important clue for joins between cones is the existence of deep concavities in the bounding contour. Figure 14a illustrates this. Provided that the end of one cone joins the side of the other well between its ends, one cannot help forming substantial concavities in the outline. The precise result that establishes this for convex cones is theorem 6, but the details may be confined to the appendix without loss. It basically states that the total concavity created by a join like that shown in figure 14a is nearly 180^0, and this gives us a method for detecting such joins. Since angles of 180^0 are preserved by linear transformations, the effect of altering the viewing angle is entirely due to changes in the angles ϕ_i that are caused by foreshortening l_1 or l_2 of figure 23. This means that the join will remain detectable until the projection of one cone's length becomes comparable with the projection of its width (when the view becomes "unconventional"), or until the junction is obscured.

A2: Cones not everywhere convex

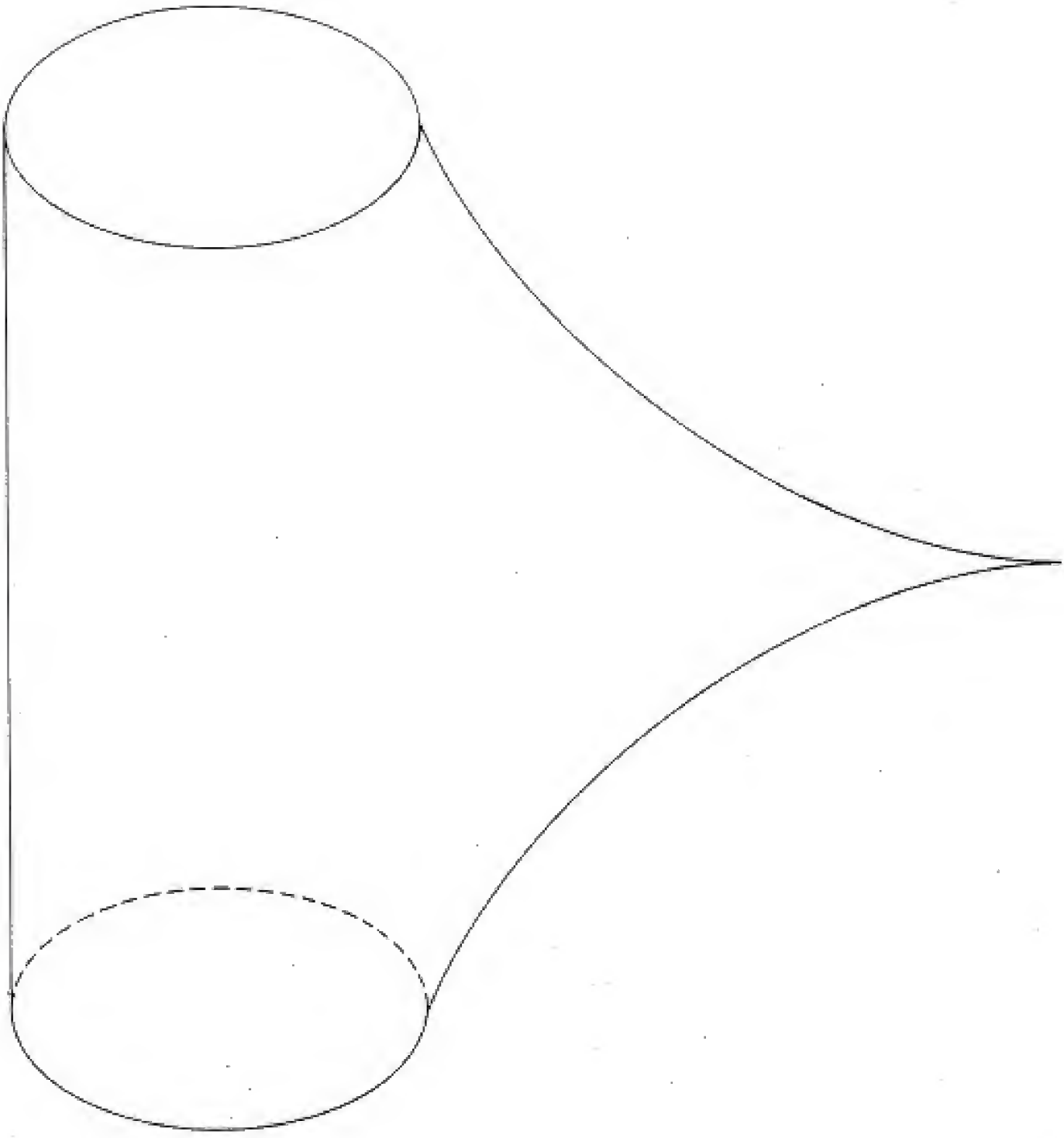
If the generalized cones Σ_1 and Σ_2 are not convex – for example, if their axial scaling functions contain concave segments – the concavities that "ought" to arise at their junction can be concealed, appearing as part of the concavities due to their axial scaling functions. The simplest case of this is shown in figure 15, where there is no identifiable concavity due to the join. Therefore, although concavities provide our algorithms with useful first places to look for joins, we need somewhat more solid results on which to base the underlying theory of join detection.

The approach we take for diagnosing joins is similar to that of theorems 3 to 6. We establish a set of constraints that are satisfied by the different types of join, and argue that, if the contours in an image decompose into segments that satisfy these constraints, then they may be interpreted as two joined generalized cones. If there is only one decomposition of the contours that satisfies these constraints, then the interpretation is also unique. The relations involved are usually quite simple. We shall assume that *R5* holds throughout this section.

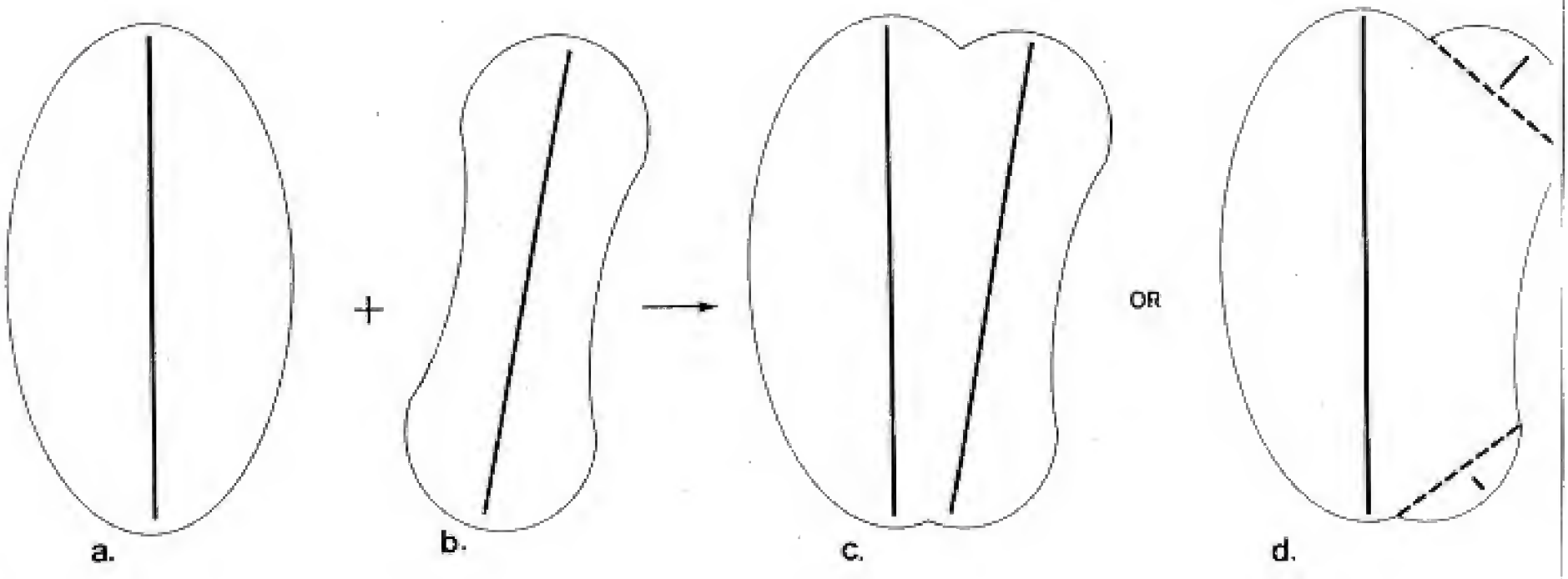
Suppose that an end of Σ_2 joins the side of Σ_1, and that the resulting surface is viewed along the direction perpendicular to the planes of the axes. If the angle ω (figure 14) between their axes is small, or if the line of sight lies too near the plane of the axes, only one "side" of each cone may be visible (figure 16). In such cases, there are no symmetry relations in the image, and the cones' axes cannot be found.

Provided that both sides of the cones remain visible (figure 17), convex and concave segments that lie distal to the join are uninfluenced by it and will obey the symmetry theorem 3. In this way, the distal segments of the cones determine their axes, which can then be extended back to the join (shown dotted in figure 17).

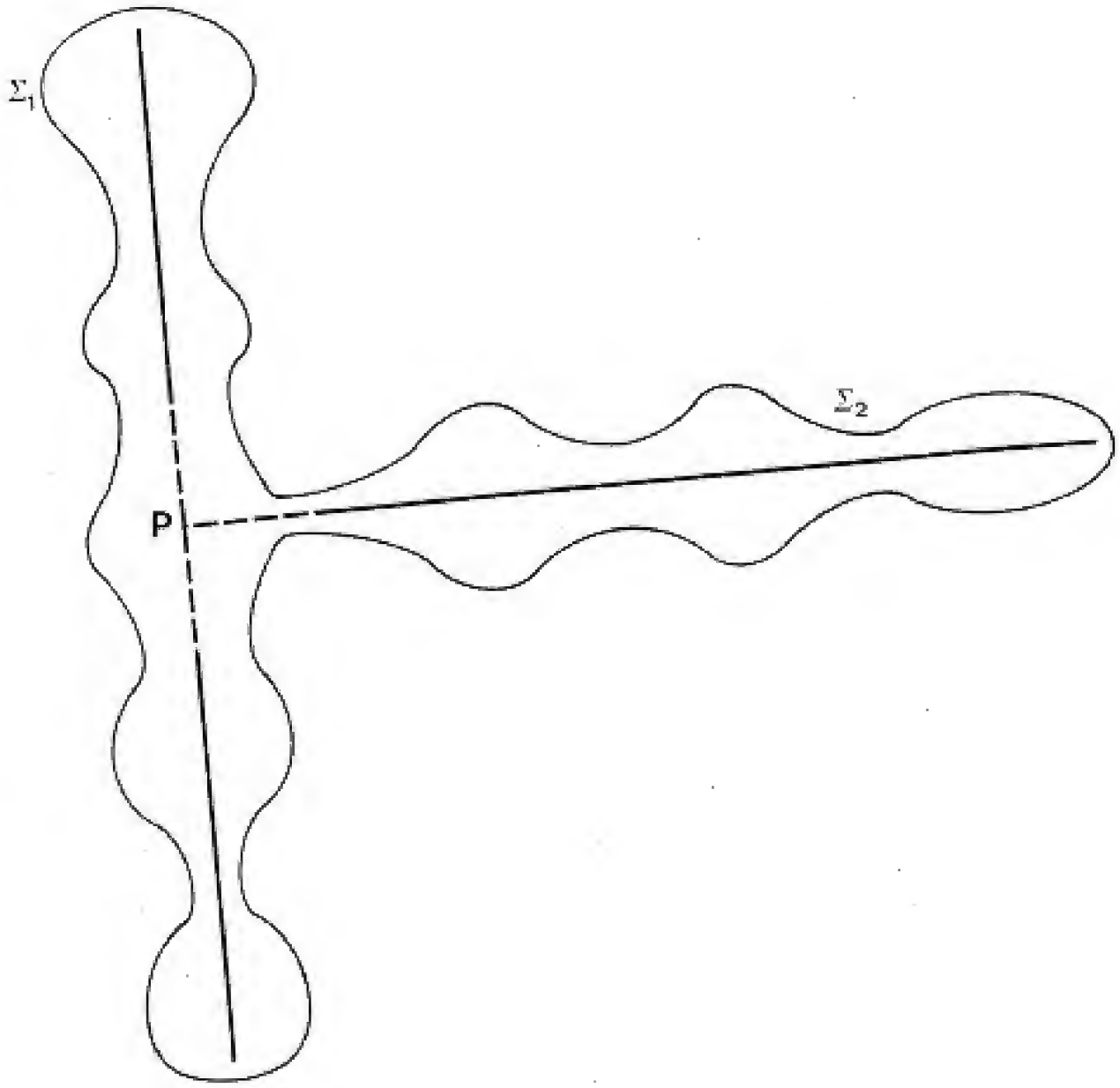
This diagnostic technique relies on the existence of segments distal to the junction, so we now deal with the case in which there are none. If we assume that the join takes place entirely within one segment of Σ_2, there are six possible situations and they appear in the top and bottom rows of figure 18. Four arise when Σ_1 and Σ_2 consist of just one segment each, and it may be either convex or concave; the other two arise because Σ_2 can straddle a segment boundary in Σ_1 (column 3 of figure 18). It is convenient to subdivide the cases



15. Joins between cones can easily be hidden. Here, concavities that "ought" to arise as a result of the junction are hidden in concavities that are caused by the axial scaling function. This is why straightforward methods for finding joins between cones fail in the general case.



16. If two cones join at so oblique an angle that the structure of one is intermixed with the structure of the other, it may be impossible to recover their axes from the image. This figure shows an example in which the cones of 16a and 16b are joined to form 16c. Methods based on the theory given here would produce the decomposition shown in 16d rather than that of 16c, because there is no detectable symmetry about the axes depicted in 16c.



17. If two joined cones are both long, which means that they both contain more than one segment distal to the join, their axes can be recovered by methods based on theorem 3 that take no account of the join. The axis fragments thus obtained, shown solid, can then be extended to their intersection point along the dotted lines, and the join itself can be analyzed after this.

where Σ_2 is concave into those cases in which the value of s_2 passes through a minimum and then increases as one moves distally from Σ_1 (bottom row), and those in which the minimum value of s_2 is achieved at the distal end of Λ_1 (middle row). In Figure 18, this minimum value is zero, which produces a cusp-like Σ_2. Notice that case 18a exhibits the situation described by theorem 6, which guarantees the presence of the concave segments joining Σ_1 and Σ_2, for reasonable values of l_1 and w_1.

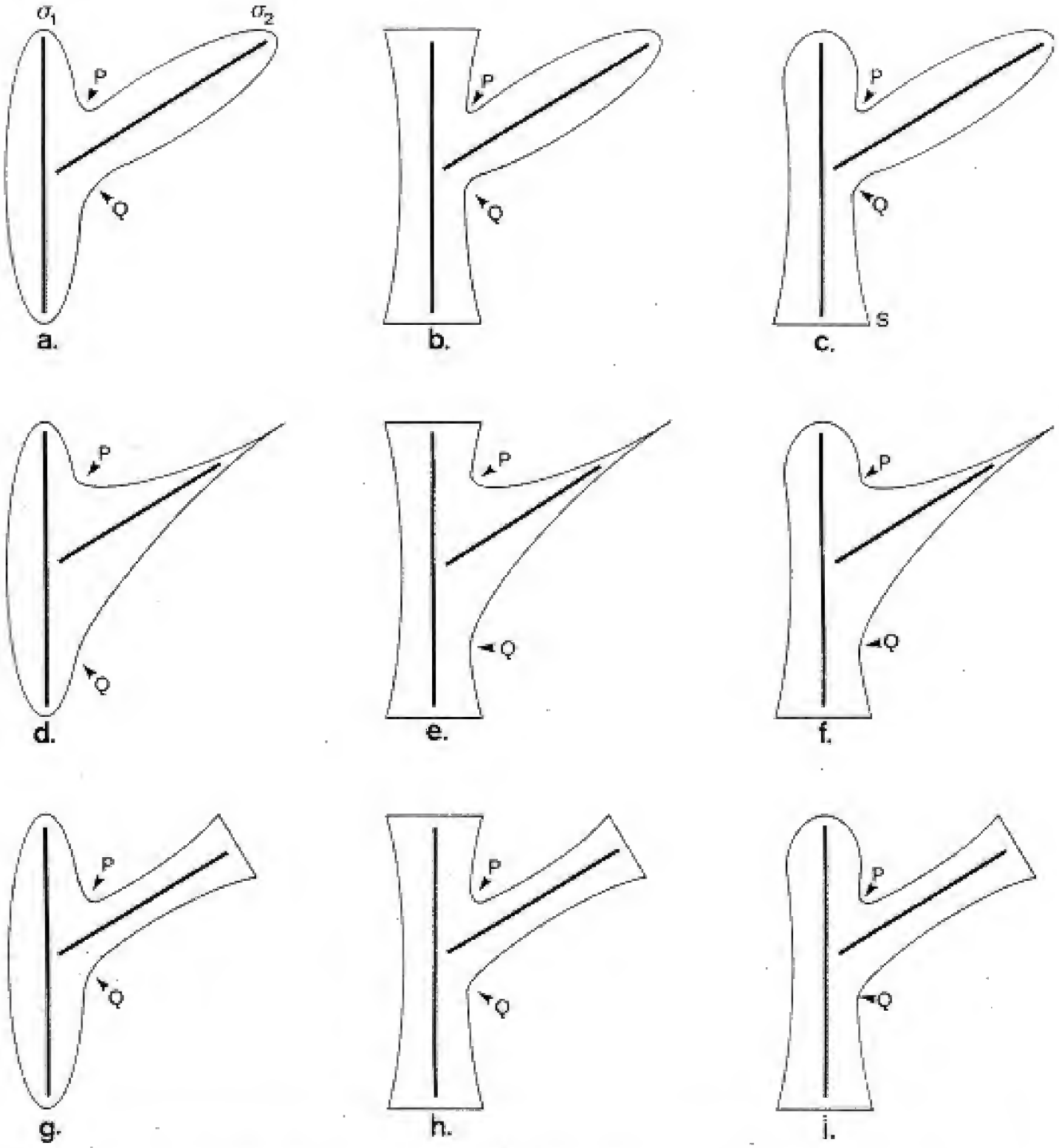
We are now ready for the main result about side-to-end joins. Figure 18 explains what is happening. In each of the cases shown there, segmentation points P and Q can be found that decompose the contour so as to satisfy a number of relationships. Theorem 7 defines the segmentation points precisely by making these relationships explicit.

> *Theorem 7 (Side-to-end joins).* Let C_V be a connected contour bounding the image of two generalized cones Σ_1 and Σ_2 connected by a side-to-end join. We assume that the image is formed from a distant viewpoint chosen such that the viewing direction lies perpendicular to the plane containing the two cones' axes Λ_1 and Λ_2. Assume that C_V is broken into segments at points of inflexion. Then there exist two points P and Q each of which is either a point of inflexion or lies within a concave contour segment such that
> (i) The line PQ lies within the figure bounded by C_V
> (ii) PQ divides C_V into two parts C_1 and C_2, between the contour segments or fragments in each of which there exists a qualitative symmetry whose two axes are the images of Λ_1 and Λ_2
> (iii) P and Q minimise the length of contour fragment left unmatched by these symmetries
> (iv) contour fragments in C_1 left unmatched by the symmetry round Λ_1 would be matched by contours whose proximal parts, and possibly all of which, lie in the interior of $C_2 \cup PQ$; and *vice versa*.
> (v) the image of Λ_2 intersects PQ between P and Q
>
> *Corollary:* If the points P and Q are unique, these constraints determine a unique decomposition of C_V from which images of the two axes Λ_i may be recovered.

In practise, it does not matter if P and Q are not unique provided that all possible choices give the same axes.

B: Two generalized cones joined end-to-end

If Σ_1 or Σ_2 contains more than one convex or concave segment, that cone's axis may be found for segments distal to the join, just as they were found in figure 17 for side-to-end joins. Hence we need consider only the case where Σ_1 and Σ_2 have just one segment. Once again, the main result depends on characterizing the segmentation points P and Q, and figure 19 gives examples of segmentation points for end-to-end joins between the various types of single segment cone. Theorem 8 defines these points precisely; it is very similar to theorem 7.



18. If the joined cones are short, the method of figure 17 cannot be used. This figure illustrates the types of side-to-end join that can occur. In the first column, the left-hand cone is convex; in the centre column it is concave, and in the third column, it is convex on one side of the join, and concave on the other. The other cone is convex in the top row, and concave in the other two. Segmentation depends upon finding the points P and Q, which are defined in the text by theorem 7 and illustrated here for each case.

Theorem 8 (End-to-end joins). Let C_V be a connected contour bounding the image of two generalized cones Σ_1 and Σ_2 connected by an end-to-end join. We assume that the image is formed from a distant viewpoint chosen such that the viewing direction lies perpendicular to the plane containing the axes of the two cones A_1 and A_2. Assume that C_V is broken into segments at points of inflexion. Then there exist two points P and Q in C_V such that:

(i) Either may be a point of inflexion, one (but only one) may lie within a concave segment, and one (but only one) may lie within a convex segment.

(ii) The line PQ lies within the figure bounded by C_V

(iii) PQ divides C_V into two parts C_1 and C_2, between the contour segments or fragments in each of which there exists a qualitative symmetry whose two axes are the images of A_1 and A_2

(iv) P and Q minimise the length of contour fragment left unmatched by these symmetries

(v) contour fragments in C_1 left unmatched by the symmetry round A_1 would be matched by a contour whose proximal parts at least lie in the interior of $C_2 \cup PQ$; and *vice versa*.

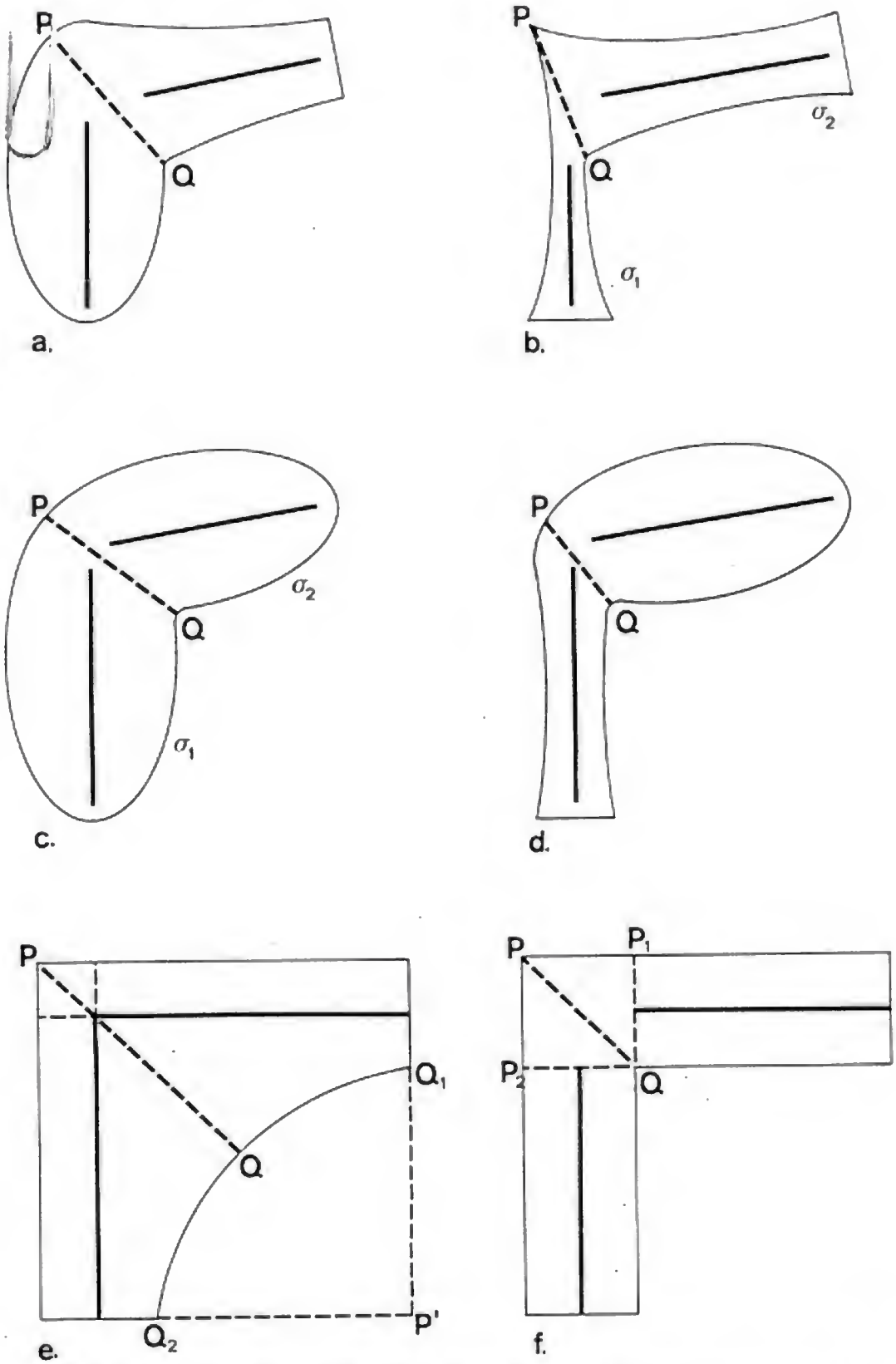
(vi) the images of A_1 and A_2 intersect PQ between P and Q

Corollary: If the points P and Q are unique, these constraints determine a unique decomposition of C_V from which images of the two axes A_1 may be recovered.

Extension to cases where some contour segments are straight lines

The assumptions *R1* - *R4* that were made about Σ_1 and Σ_2, excluded cases where these surfaces contained straight lines. Such cases are frequent in real life, however, and some examples are shown in figures 19e and f. 19e is a limit of 19d, and in some sense also of 19a; 19f is a limit of all of the cases. 19f may be solved in the standard way; Q is the only concave point in the contour, and it matches either the point P, or it induces two "nearest" points P_1 and P_2 that separate the two arms of the figure from the rectangle QP_1PP_2. Both segmentations are permissible.

Case 19e is more difficult. The only true inflexion points are Q_1 and Q_2, but the line Q_1Q_2 lies outside the figure. If Q_1 and Q_2 are used despite this, the segmentation to which they lead corresponds to thinking of the figure as a rectangle with a piece excised (cf Hollerbach 1975 p. 55). This would be the preferred description if Q_1 and Q_2 lie near P'. Since 19e may be regarded as the limit of 19d, the point P (a corner joining two staight lines) can be regarded as a segmentation point, like the point P in 19d. P then induces the point Q as shown, which segments the figure in the same way as 19d. When designing algorithms for dealing with cases where some lines are nearly straight, "convex" corners often acquire a dual status that arises from regarding the straight lines as limits of concave rather than convex contour segments. This means in practise that straight lines are somewhat more difficult to deal with than curves since, in the initial state of the algorithms for implementing the methods defined here, straight lines and the corners to which they lead



19. 19a to d exhibits the possible types of end-to-end joins, for cones that contain only one segment. The segmentation points P and Q are defined by theorem 8, and illustrated here for the different cases. They provide the basis for interpreting the join from an image. 19e & f exhibit the configurations that arise in limiting cases where some or all of the contours become straight lines. Notice that in 19f, the symmetry relations have degenerated into

may be associated with several possible labellings.

C: Joins between more than two generalized cones

The principal difference between this and case B above is that a given point of segmentation may have more than one match elsewhere. For example, in the silhouette of an octopus, the deep concavity between each tentacle matches two others. Also in this case, it is possible to have end-to-end joins in which both P and Q lie within concave contour segments. The only straightforward result about the case of multiple joins holds when all the joined axes are coplanar, which is a common but restrictive condition. In this case, the relevant result is so similar to theorems 7 and 8 that I omit it.

4: Discussion

The purpose of this article was to elucidate the assumptions that can reasonably be made when interpreting the occluding contours in an image. The assumptions at which we arrived were stated as restrictions *R1* - *R4*, and it was then proved that these restrictions have a close relationship to the assumption that the viewed surface is composed of generalized cones. In the second and third parts of the article, we took this result as an assumption and studied properties of images of surfaces constructed in this way. We found that many constraints hold among portions of the contours in images of such surfaces, and that rough symmetries are formed around the image of the cones' axes. The importance of these relations is that one can use them to design algorithms for finding the generalized cone-based description of a contour, and for extracting any axes that may be present. By applying these algorithms repeatedly to the contours found in an image, one can often derive the 3-D model representation (Marr & Nishihara 1977) of a surface's shape without prior knowledge of it. Methods based on the present theory will however fail for views in which one or more axes are foreshortened (roughly, whenever the condition $l.sin(\chi) > d$ of page 17 is violated).

The theory presented in this article is a pure competence theory, or a theory at the topmost of Marr & Poggio's (1976) four levels. It is concerned with ends not means. The natural division between means and ends is interestingly illustrated by the methods for segmenting a contour into two component generalized cones (theorems 7 & 8). The starting-point for our algorithms that actually find the points P and Q as defined by these theorems is the examination of deep concavities in the contour C_V (see Vatan 1976). This contrasts strongly with the theory, because the concavities may be small or even absent, especially for end-to-end joins. Only in certain circumstances does the underlying theory guarantee their presence (theorem 6).

Acknowledgements: I thank Harold Abelson, Keith Nishihara and Shimon Ullman for careful readings of early drafts of the manuscript, and for suggesting improvements to several results. Karen Prendergast prepared the figures. This report describes research

done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643.

References

Binford, T. O. 1971 Visual perception by computer. Presented to the IEEE Conference on Systems and Control, Miami, in December 1971.

Hollerbach, J. M. 1975 Hierarchical shape description of objects by selection and modification of prototypes. *M. I. T. A. I. Lab. Technical Report 346.*

Marr, D. & Nishihara, H. K. 1977 Representation and recognition of the spatial organization of three dimensional shapes. (Submitted for publication).

Marr, D. & Poggio, T. 1976 From understanding computation to understanding neural circuitry. In *The visual field: psychophysics and neurophysiology. Neurosciences Research Program Bulletin*, E. Poeppel *et al.*, Eds. (in the press).

Vatan, P. 1976 Segmentation of figure after separation from ground. *M. I. T. Bachelor's Thesis Report.*

Vatan, P. & Marr, D. 1977 Algorithms for the segmentation of a contour. (In preparation).

Waltz, D. L. 1975 Understanding line drawings of scenes with shadows. In: *The psychology of computer vision*, Ed. P. H. Winston, pp19-91. New York: McGraw-Hill.

Appendix

Suppose that Σ is an arbitrary three-dimensional surface, and that S_V is its image from viewpoint V as produced by the projection ψ: $\Sigma \longrightarrow S_V$ (figure 2 in the main text). S_V has a bounding contour C_V say, that corresponds to the silhouette of Σ, and which we can think of as having been obtained by acting on S_V with a boundary operator ∂ (see figure 2).

Definition. Let Γ_V be the set of points on Σ whose image lies on C_V.

Then Γ_V is called the *contour generator* of C_V on Σ.

That is, Γ_V is the set of points P on Σ such that $\psi(P)$ lies on C_V. We can now define the operator δ: $\delta(\Sigma) = \Gamma_V$ which is induced by ∂, and which selects Γ_V out of Σ. This is illustrated in figure 2, where ∂ is defined is such a way that the diagram in figure 2 commutes — i.e. $\psi\delta = \partial\psi$. Notice that ψ, C_V, S_V and Γ_V all depend upon the vantage point V.

A formal statement of the restrictions *R1* - *R5* is now given:

Restriction R1: Σ is everywhere twice differentiable with continuous second derivative.

Restriction R2: The inverse ψ^{-1} : $C_V \longrightarrow \Gamma_V$ is one-valued.

Restriction R3: The mapping ψ : $\Gamma_V \longrightarrow C_V$ is continuous.

Restriction R4: The contour generator Γ_V of C_V is planar.

Restriction R5: The axes of two joined generalized cones are coplanar.

Lemma 1. Let $f(x, y) = 0$ describe a planar curve that is twice differentiable

with continuous second derivative. Let $L : R^2 \longrightarrow R^2$ be a non-singular linear transform of the plane. Then L preserves points of inflexion in f.

Proof: We define $g = Lf$, the image of f under L, by $g(x, y) = f(L^{-1}(x, y))$. Since L is linear, non-singular, and therefore continuous, it induces a *(1 - 1)* correspondence between the slopes of tangents to f and tangents to g. We can represent the set of possible slopes by the unit circle S^1, and so L induces a map $L^\circ : S^1 \longrightarrow S^1$, which is *(1 - 1)* and onto. Hence L° is monotonic (either increasing or decreasing) – that is, if θ_3 lies between θ_1 and θ_2, and $|\theta_1 - \theta_2| < 2\pi$, then $L^\circ(\theta_3)$ lies between $L^\circ(\theta_1)$ and $L^\circ(\theta_2)$. Now a point of inflexion X on f is a stationary point for the slope of the tangents to f. Because L° is monotonic, $L(X)$ is therefore a stationary point for the slope of the tangents to $Lf = g$. Hence $L(X)$ is an inflexion point on g.

I have used a geometric rather than an analytic argument because it is clear how the same argument applies to the case where f is piecewise linear. In this situation, the analog of an inflexion point is a point where the sign of the change in gradient reverses, and the argument used here still applies.

Definition. Let $\rho(r, \theta) = 0$ be a simple closed planar curve that is twice continuously differentiable; and let h be a twice continuously differentiable positive real function. Let Λ be a line at some angle ψ to the plane of ρ, and denote positions along Λ by the variable z. Let Σ be the surface $h \times \rho$. Then Σ is a *generalized cone* with *axis* Λ, *cross-section* ρ, *scaling function* h.

and *eccentricity* ψ. If $\psi = \pi/2$, Σ will be called a *right generalized cone*. (See figure 5 in the main text).

Definition. Let V be a distant vantage point for the generalized cone Σ such that (i) the image formed from V is an orthogonal projection, and (ii) the rays in the projection all lie parallel to the plane of the cross-section of Σ. Let the direction of these rays in the plane be denoted by the angle ϕ. When these restrictions are in effect, we shall denote the contour generator Γ_V by Γ_ϕ.

Theorem 1. Let Σ be a generalized cone with convex cross-section $r = \rho(\theta)$. Then Σ satisfies *R1* everywhere, and for all orthogonal projections parallel to the cross-section ρ, it satisfies the conditions *R2* - *R4*. Conversely, if *R1* is satisfied by the closed surface Σ, and if *R2* - *R4* are satisfied for all orthogonal projections parallel to some plane Π, then Σ is a generalized cone with convex generating cross-section ρ that lies parallel to Π.

Proof: Our definition of a generalized cone ensures that it satisfies *R1*. Since Σ is generated by moving ρ along the axis Λ (see figure 5), a given radial $PG = (\rho(\theta), \theta)$ sweeps out a plane that contains Λ, as ρ itself is moved along Λ. As ρ moves, the radial PG maintains its direction, but shrinks or expands in a manner dictated by the scaling function $k(z)$. As G moves, it traces out a curve on Σ, which we shall call Γ_V, and which lies in the plane Λ.

Furthermore, the tangent to Σ at G that lies in the plane of ρ is the tangent GV to ρ at G, for all positions of G. Suppose that we represent the direction of GV in the plane of ρ by ϕ. If one views Σ from a great distance in the direction ϕ, the line GV is a line of sight to the edge of the surface Σ. Therefore G lies on the contour generator for this view of Σ. But this is true for all positions of PG as P moves along Λ, and so Γ_ψ is a contour generator. In fact it is Γ_ϕ. Furthermore, since ρ is convex, the same will be true for every angle ϕ and corresponding point G on ρ, provided that the viewing directions lie parallel to the plane of ρ. Hence Σ satisfies *R2* - *R4* for all such orthogonal projections.

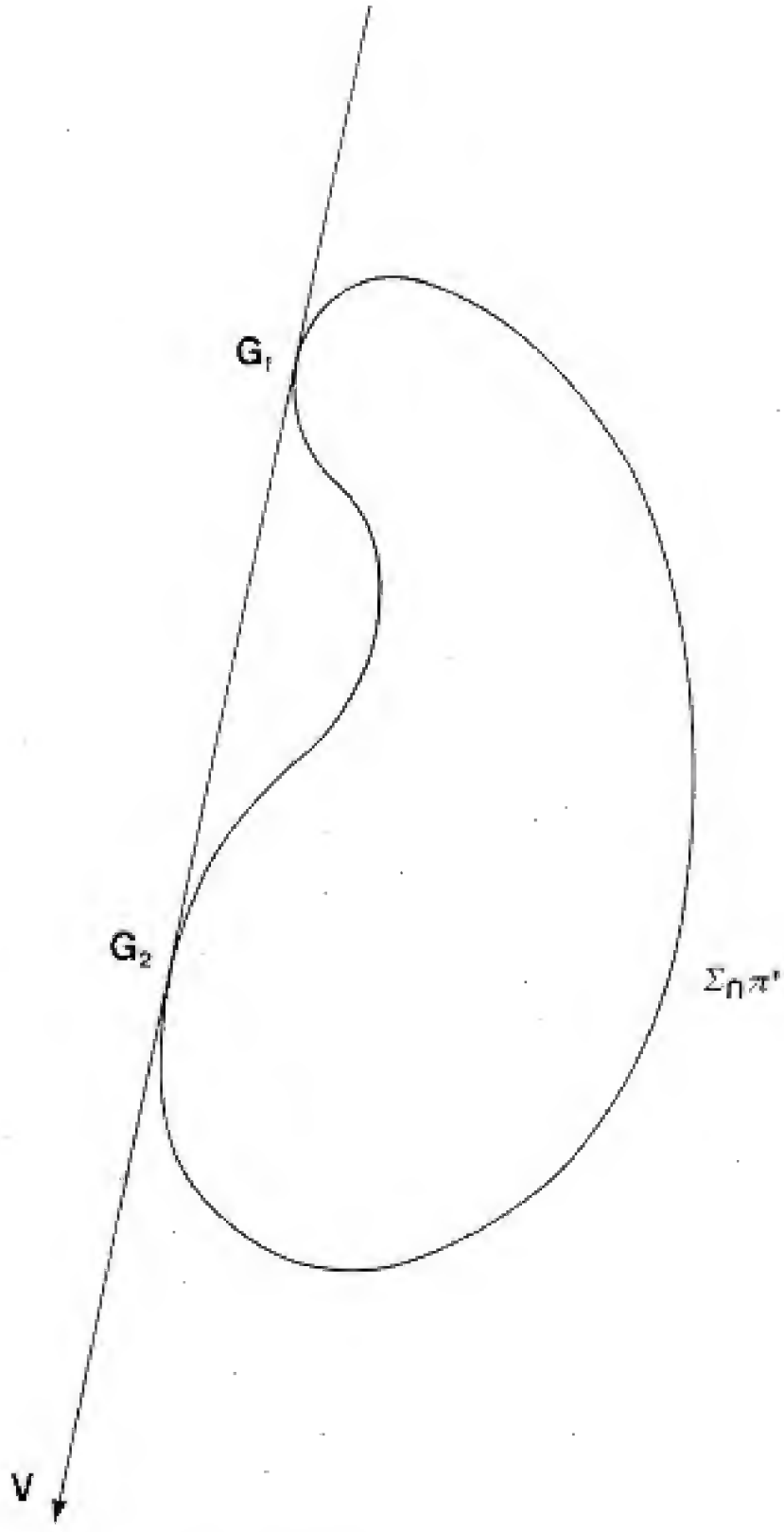
The proof of the converse result is longer, and we first need to establish three lemmas.

Lemma 2. $\Sigma \cap \Pi'$ is convex for all planes Π' parallel to the plane Π of the given viewing directions.

Proof: Suppose that $\Sigma \cap \Pi'$ were not convex. Then there would exist a line in Π' that was tangential to $\Sigma \cap \Pi'$ at two points G_1 and G_2 say, as shown in figure 20. But the line G_1G_2 is the ray that produces the edge of the image of Σ from this viewing angle, and G_1G_2 therefore projects to a point P say, on C_ψ. So $\psi^{-1}(P)$ would contain both G_1 and G_2, and so would not be single-valued. This contradicts *R2*.

Lemma 3. If two distinct contour generators on Σ intersect at a point X, then contour generators for all distant viewing directions in the plane Π pass through X.

Proof: The tangent plane to Σ at X, which exists by *R1*, contains two distinct vectors that lie



20. Lemma 2 shows how restriction $R2$ forces ρ to be convex.

in a plane parallel to Π. Hence the tangent plane at X must itself be parallel to Π.

Lemma 4. Let Γ_{ϕ_1} and Γ_{ϕ_2} be contour generators for two different viewing directions in Π. Then Γ_{ϕ_1} and Γ_{ϕ_2} intersect on Σ.

Proof. Since Σ is a closed surface, the Γ_ϕ for any angle ϕ divides Σ into two components. This follows from the fact that if Σ is the surface defined by the equation $f(x, y, z) = 0$, the points on Γ_ϕ are solutions to the equations

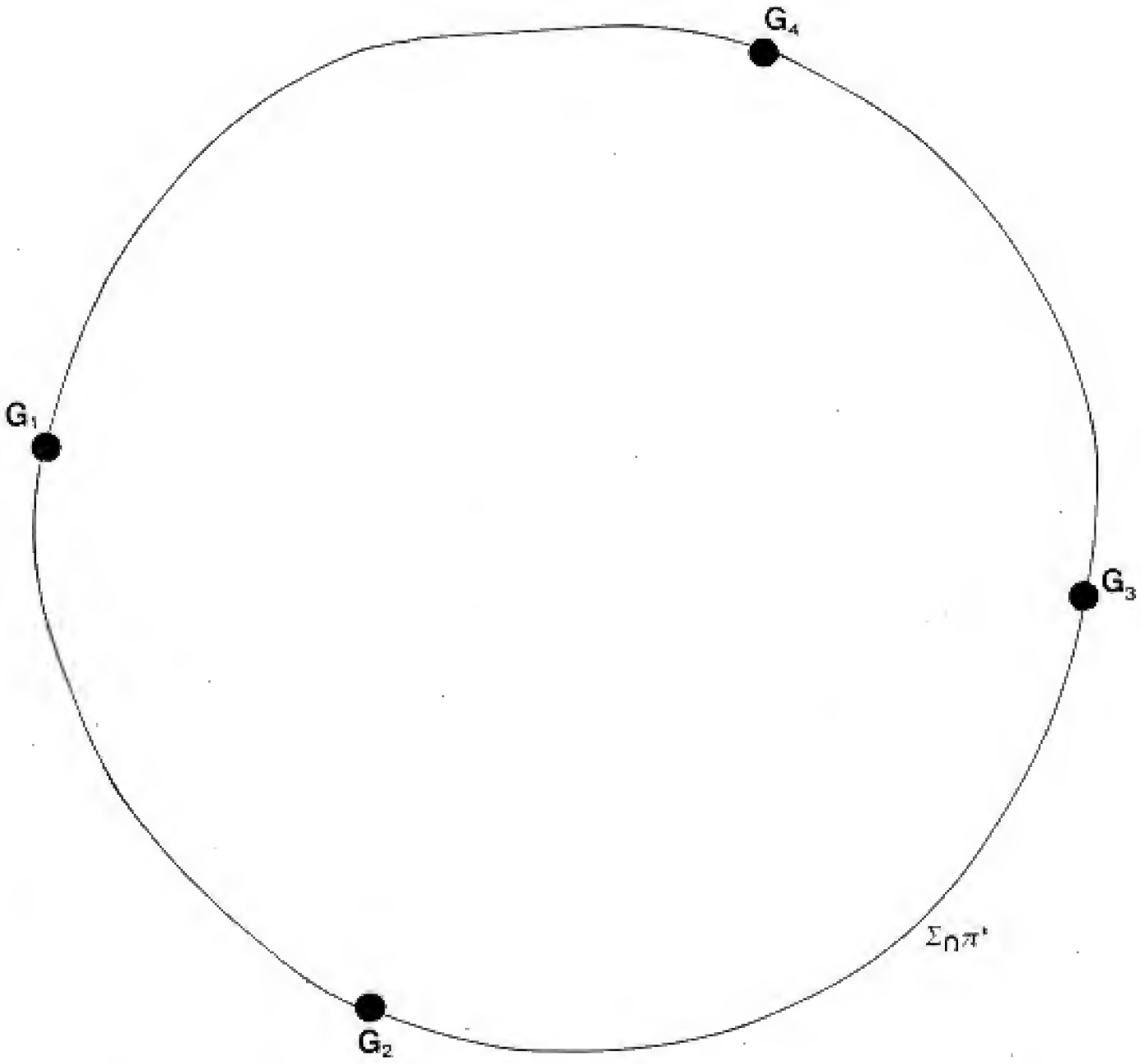
$$Grad(f) \cdot (V - (x, y, z)) = 0 \qquad (1)$$

$$f(x, y, z) = 0 \qquad (2)$$

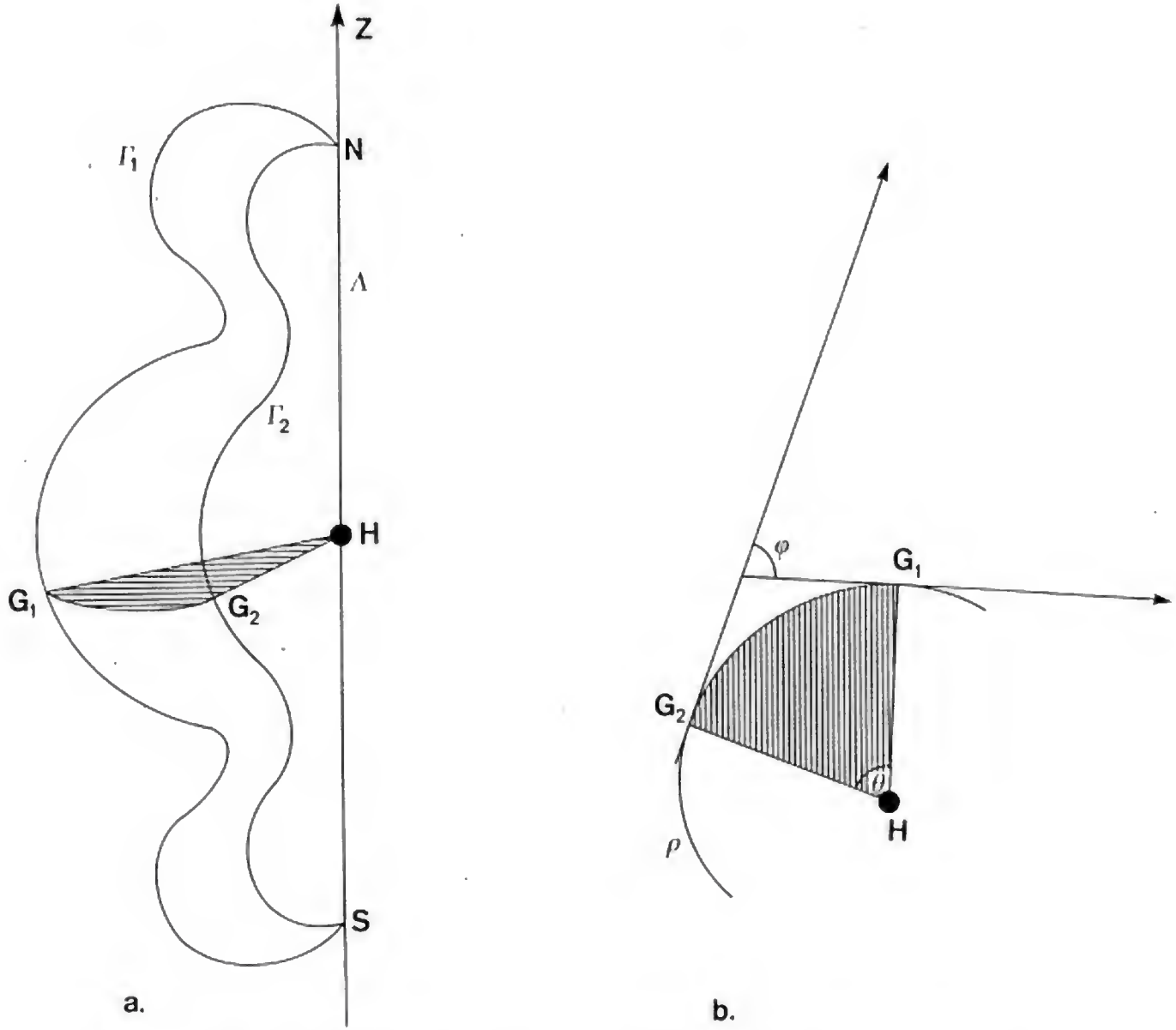
where V is a distant vantage point along the rays of the orthogonal projection Γ_ϕ, and the two components correspond to points where equation (1) takes values > 0 and < 0 respectively. Hence Γ_{ϕ_1} and Γ_{ϕ_2} each divide Σ into two connected components. Let Π' be any plane parallel to Π that intersects Σ in more than a point. $\Sigma \cap \Pi'$ is a simple closed convex curve, which meets Γ_{ϕ_1} in G_1 and G_3, and which meets Γ_{ϕ_2} in G_2 and G_4, as illustrated in figure 21. The tangents to Σ at G_1 and G_3 are parallel, and so are the tangents at G_2 and G_4. Clearly, the line G_1G_3 divides the simple closed curve $\Sigma \cap \Pi'$ into two parts, in one of which lies G_2, and in the other of which lies G_4. But G_2 and G_4 both lie on Γ_{ϕ_2}, whereas G_1 and G_3 both lie on Γ_{ϕ_1}. Hence Γ_{ϕ_1} and Γ_{ϕ_2} must intersect somewhere on Σ.

Corollary: Γ_{ϕ_1} and Γ_{ϕ_2} intersect twice or more.

We can now complete the proof of theorem 1. Let Γ_{ϕ_1} and Γ_{ϕ_2} be two contour generators for Σ for different viewing directions lying in Π. Since Γ_{ϕ_1} and Γ_{ϕ_2} are both planar (by *R4*), their containing planes intersect in a line Λ (say). By lemma 4, Γ_{ϕ_1} and



21. Lemma 4 establishes that contour generators for two different viewing angles must intersect.



22. Diagrams for the proof of theorem 1. Γ_1 and Γ_2 are two planar contour generators whose planes intersect in the line NS. The shaded region is parallel to the plane of the viewing directions. Figure 22b shows a view of this region from above.

Γ_{ϕ_2} intersect in at least two points, and these points must therefore lie along the line Λ. Let N be a boundary point of the set of intersection points of Γ_{ϕ_1} and Γ_{ϕ_2} on Λ, and let S be the next closest such point to N. This situation is depicted in figure 22. By lemma 2, all contour generators pass through N and S, which we may therefore think of as north and south poles of Σ, and therefore the planes of all countour generators for views of Σ from Π must contain N and S and hence the line Λ. That is, the planes of all contour generators for distant views from the plane Π will intersect in Λ.

Let Π' be a plane parallel to Π lying between N and S, distant z from N, as shown in figure 22a. Let Π' intersect Λ at H, Γ_{ϕ_1} at G_1, and Γ_{ϕ_2} at G_2. The configuration in Π' is shown in figure 22b. The crucial step in the proof is to notice that up to scalar magnification, the geometry of figure 22b is independent of z, the position of Π' along the line NS. This follows from the following observations:

(i) The angle between HG_1 and HG_2 is independent of z, because it is simply the angle between the planes of Γ_{ϕ_1} and Γ_{ϕ_2} measured parallel to Π;

(ii) The direction of the tangent to ρ at G_1 is independent of z, because as z increases, G_1 traces out the contour generator Γ_{ϕ_1}, which is by definition the locus of tangents to Σ parallel to Π for a given fixed viewing direction ϕ_1.

We deduce that for each angle θ in figure 22b, the tangent to the curve ρ has a constant direction for every z. That is, for each z, the cross-sections ρ of Σ in Π' are all solutions of the *same* equation

$$r.d\theta/dr = f(\phi) \qquad (1)$$

where f is some function of the viewing angle ϕ and is independent of z. Let $R(\theta)$ be a

solution to *(1)*. Then the cross-section function ρ of Σ has the form

$$\rho(\theta, z) = h(z).R(\theta) \qquad (2)$$

where h is a positive real function of z. Finally, Σ is twice continuously differentiable, by *R1*. Hence h is a twice continuously differentiable function of z, and so is R of θ. This completes the proof of theorem 1.

Corollary 1: $h = 0$ at N and at S.

Corollary 2: The cross-section ρ can change at the poles where $h = 0$.

Theorem 2. A necessary and sufficient condition for Σ to satisfy *R1*, and

R2 - R4 for all distant viewing positions, is that Σ be a quadratic surface.

Proof: If Σ is a quadratic surface, it satisfies *R1 - R4*. This follows from the following three observations:

(i) A sphere trivially satisfies *R1 - R4*

(ii) Any linear transformation or translation preserves *R1 - R4*

(iii) Any quadratic surface may be derived from a sphere by a linear transformation and a translation.

Conversely, suppose that Σ satisfies *R1 - R4* for all viewing angles. Since the conditions of theorem 1 hold for every viewing angle, Σ may be thought of as a generalized cone with generating cross-sections in any direction. Hence, any two parallel planes intersect Σ in curves that, if not null, have the same shape. Suppose that Σ is the surface represented by the polynomial $f(x, y, z) = 0$. Then a set of such parallel planes is given by the family $z = ax + by + c$, for varying c. We deduce that the curves

$$f(x, y, ax + by + c_1) = 0$$

$$f(x, y, ax + by + c_2) = 0$$

are identical, up to magnification and a translation. Hence the C_i cannot multiply terms of second or higher order. Therefore, in the original equation, z cannot multiply terms of second or higher order. Since the condition holds for arbitrary planes, it must also be true of x and of y, and it must hold identically. Hence f is at most quadratic, and we have already seen that quadratic surfaces satisfy the conditions of the theorem.

Remarks about the proofs of theorems 1 and 2

The premises of these theorems, and the way the premises are used in the proofs, seem generally reasonable with one possible exception. That is the use of *R2* to show that $\Sigma \cap \Pi'$ is convex in lemma 2. In fact, this is how *R2* forces the cross-section of Σ to be convex, which is a condition that holds all through these results. One might argue that this is a somewhat artificial use of *R2*, which was introduced mainly to exclude surfaces that contain lines. The analysis given so far concerns only the silhouette of the image of Σ however, and in these circumstances the convexity assumption for ρ is a reasonable one, because violations cannot be detected from viewing directions that lie parallel to the plane of ρ. In practise, one can apply the above analysis to all occluding contours in an image, provided that one subsequently relates the different cylinder descriptions that emerge for different parts of the same surface. Figure 8 shows a simple example of this.

Theorem 3 (Axial Symmetry). Let $\Sigma = \rho \times h$ be a generalized cone with

convex cross-section ρ, viewed distantly from its associated viewing plane Π, and let the cross-section scaling function $h(z)$ contain at least one concavity. Then

(i) the silhouette of Σ decomposes into $n > 2$ contour segments by splitting it at points of inflexion;

(ii) the image of the axis Λ of Σ establishes an axial symmetry between at least 2 or $(n - 2)$ (whichever is greater) contour segments, including all concave segments, such that segments that correspond under the symmetry are either both convex or both concave;

(iii) if c_{i1} and c_{i2} are corresponding segments, and if $d(c_i)$ denotes the average perpendicular distance from c_i to the image of the axis, then $d(c_{i1})/d(c_{i2})$ is independent of i.

Proof: Since Σ is a generalized cone and ρ is convex, the contour generators for its image are the curves $(h(z).\rho(\theta_1), \theta_1, z)$ and $(h(z).\rho(\theta_2), \theta_2, z)$ for fixed θ_1 and θ_2 and variable z, as shown in figure 7. Hence as $h(z)$ increases, the distance from contour generator to the image Λ° of the axis Λ increases on both sides of Λ°. Hence if $h(z)$ has a concave portion, it will generate two concave segments c_1 and c_2 in C_ϕ, one either side of Λ° and symmetrical about Λ° (with a possible lateral displacement unless Σ is a right generalized cone). Also, $d(c_1)/d(c_2)$ depends upon the value of $\rho(\theta_1)/\rho(\theta_2)$, and not on the particular choice of contour segment (i.e. of z).

These remarks hold for all convex and concave segments of h, except possibly for those at each end of Σ (see figure 7). In figure 7, there is effectively only one end segment,

namely c_3, but in general there may be 2. Hence the symmetry established by the axis Λ^* holds for at least $(n - 2)$ of C_Φ's contour segments, (or 2 of them if $n = 3$), and it matches up all contour segments that are concave, since the end segments must be convex. Finally, there must be at least three contour segments, since the premises of the theorem require that there be 2 concave ones, and C_Φ is a simple closed curve.

Definition. The type of symmetry established by theorem 3, which holds between convex or concave segments of a contour and which includes a scaling factor, we shall call a *qualitative symmetry.*

Definition. The set $\mathbf{T}$ of points $h(z_i).\rho(\theta)$ for all θ and each value of z_i that makes $h(z_i)$ a local maximum or minimum, is called the *radial extremity* of Σ. Let V be an arbitrary distant vantage point, and let V' be the projection from V onto a cross-section plane of Σ, as described in the main text and illustrated in figure 9. The set $\Gamma_{V'} \cup \mathbf{T}$ is called the *skeleton* of Σ for the vantage point V.

Theorem 4 (Skeleton Theorem). Let $\Gamma_{V'} \cup \mathbf{T}$ be the skeleton of Σ associated with some distant vantage point V. Then provided that C_V is the projection of $\Gamma_{V'}$ to V,

(i) C_V is qualitatively symmetric about the image of the axis Λ of Σ, in the sense of theorem 3

(ii) the image of $\mathbf{T}$ consists of one or more connected components, through which $\mathbf{A}$ passes, and between any two of which there exists a mapping that is *(1 - 1)*, continuous and onto, that preserves the gradient of the image of $\mathbf{T}$ at each point.

Proof: $\Gamma_{V'}$ correponds to a viewing direction that is coplanar with the cross-section ρ. Hence the contours $C_{V'}$ in the image as seen from V' obey the conditions of theorem 3, and the image of the axis $\mathbf{A}$ induces a qualitative symmetry between its concave and convex components. Such relations are preserved by an orthogonal projection, and since it is a condition of the theorem that C_V coincides with the projection of $\Gamma_{V'}$, it follows that C_V is also qualitatively symmetric about the image of $\mathbf{A}$. Secondly, $\mathbf{T}$ trivially consists of one or more connected components through which $\mathbf{A}$ passes, since these components are just cross-sections at the local maxima and minima of h. Finally, if $h(z_1)$ and $h(z_2)$ are local maxima or minima of h, the mapping between points on Σ given by

$$(h(z_1).\rho(\theta),\ \theta,\ z_1) \longrightarrow (h(z_2).\rho(\theta),\ \theta,\ z_2)$$

is continuous, *(1 - 1)*, onto, and preserves the gradient of $\mathbf{T}$ at each point, since the gradient of both at θ is $\rho.d\rho/d\theta$. This correspondence is preserved by an orthogonal projection, and hence the relations will still hold in the image of $\mathbf{T}$.

Definition. Let Σ be a generalized cone, and let Φ be the set of points $(h(z).\rho(\theta_i),\ \theta_i,\ z)$ for all z, for each θ_i that makes ρ a local maximum or minimum. Then Φ is called the *fluting* of Σ. For a viewpoint V, let $\Gamma_{V'} \cup \mathbf{T}$ be the skeleton of Σ that occurs in theorem 4. We define the *complete*

skeleton by adding the fluting to Σ's skeleton, i.e. the complete skeleton of Σ from viewpoint V is the set $\Gamma_V{}' \cup \Upsilon \cup \Phi$.

Theorem 5. Let $\Gamma_V{}' \cup \Upsilon \cup \Phi$ be the complete skeleton of Σ associated with some distant vantage point V. Then

(i) C_V and the image of Υ obey theorem 4

(ii) The image of each portion $\{(h(z).\rho(\theta_1), \theta_1, z)$, for fixed θ_1 and varying $z\}$ of the fluting of Σ is either a straight line, or it divides into convex and concave segments that are in *(1 - 1)* correspondence with the convexities and concavities in that part of C_V which lies on the same side of its axis of symmetry.

Proof: Part (i) follows from theorem 4. Part (ii) follows because like the two qualitatively symmetric components of C_V, each contour in the fluting of Σ is generated by variations in $h(x)$. If such a contour lies directly on the line of sight to the axis Λ of Σ, it will appear in the image as a straight line. Otherwise, its concavities and convexities will follow those of one component of C_V, although the depth of the concavities or convexities will differ in general.

Definition. For $i = 1, 2$ let Σ_i be a generalized cone with maximum width $2w_i$ and axis Λ_i of length l_i. Let Σ_1 and Σ_2 be joined, and let ω be the angle between their axes (see figures 14 and 23). Then the join between the cones will be called *side-to-end* (the side of Σ_1 to the end of Σ_2) provided

that

(i) the two axes intersect between the ends of Λ_1

(ii) the whole of the joined end AB of Σ_2 lies between the two lines perpendicular to and passing through the ends of Λ_1

(iii) one of the points A and B of figure 14 does not lie within the convex hull of Σ_1.

The join is called *end-to-end* if

(i) the two nearest ends of Λ_1 and Λ_2 are within $min\ (w_1, w_2)$ of one another, and

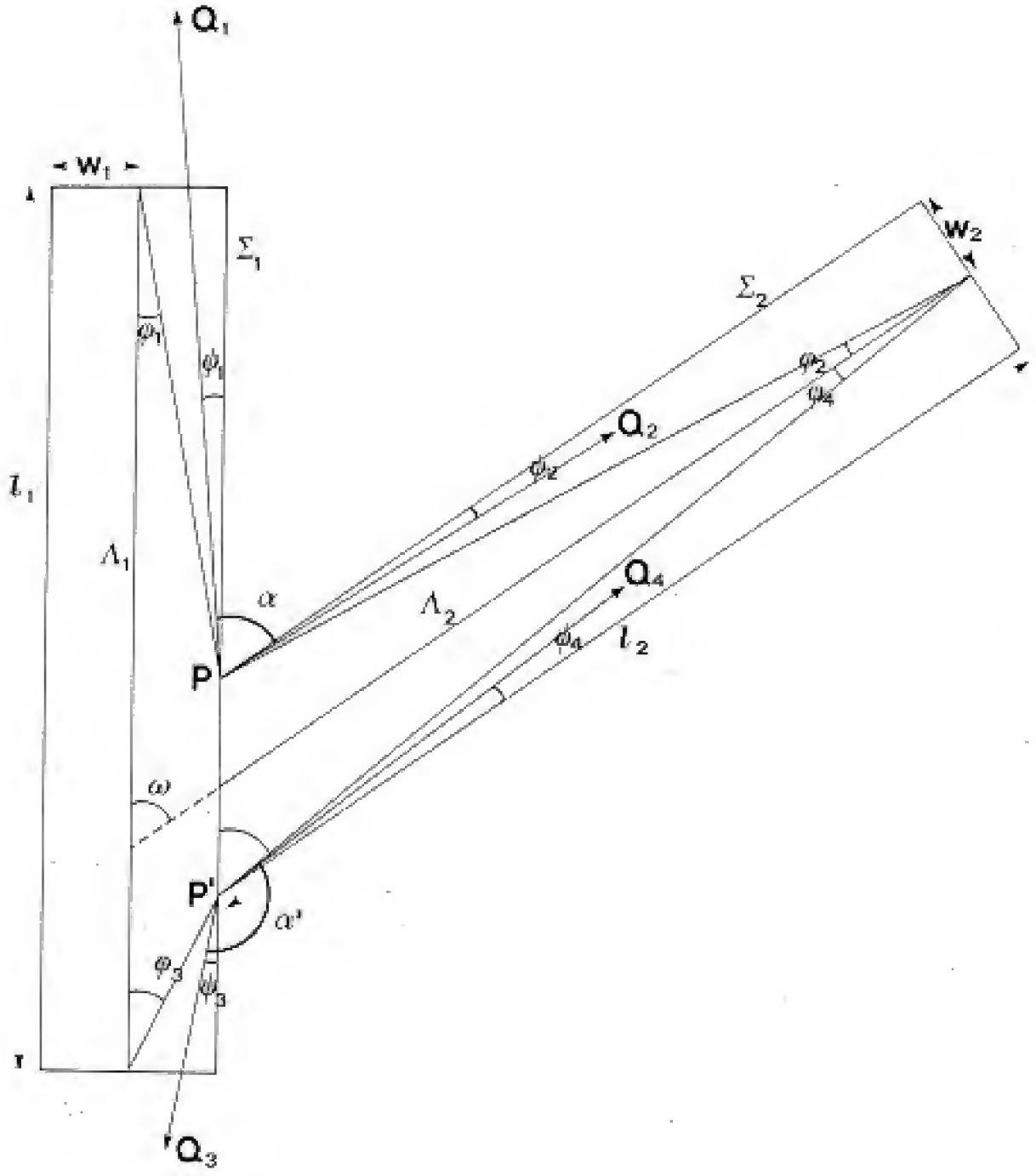
(ii) the two furthest ends are greater than $(w_1 + w_2)$ apart.

Theorem 6. Let Σ_1 and Σ_2 be two convex generalized cones such that the end of Σ_2 joins the side of Σ_1, and the join satisfies restriction *R5*. Let the lengths of the cones' axes Λ_i be l_i, and let the diameters of their cross-sections be bounded by $2w_i$ $(i = 1, 2)$. Then

(i) the only concavities that can occur in the image are due to the junction

(ii) viewed distantly perpendicular to the plane of Λ_1 and Λ_2, the total concave angle present in the image is near 180^0 (in the sense made precise in the proof) provided that the site of the join is not near an end of Σ_1, and that Σ_1 and Σ_2 are much longer than they are wide.

Proof: By lemma 5, all contours derived separately from Σ_1 and Σ_2 are convex. Hence any concavity in the image of their union must be due to the way they are joined. If Σ_1 and Σ_2



23. Diagram for the proof of theorem 6. The idea is to obtain lower bounds for the concavities in the outline that are due to the join. The total concavity is $(\alpha + \alpha')$, which is near 180^0.

have coplanar axes and are viewed perpendicularly to this plane, the resulting configuration is as represented in figure 23. The contours shown in thick lines there represent cylinders l_i long and w_i thick which, by the conditions of the theorem, bound the cones Σ_i $(i = 1, 2)$. Let P be one of the two points at which contours due to Σ_1 and Σ_2 intersect, and let PQ_1 and PQ_2 be the tangents to Σ_1 and Σ_2 at P. Let ω be the angle between the axes Λ_1 and Λ_2 of Σ_1 and Σ_2, and let ψ_1 and ψ_2 be the angles that PQ_1 and PQ_2 make with Λ_1 and Λ_2. Then the angle between PQ_1 and PQ_2 is $\alpha = (180^0 - \omega - \psi_1 - \psi_2)$. The corresponding angle at the other intersection P' between contours of Σ_1 and Σ_2 is $\alpha' = (\omega - \psi_3 - \psi_4)$, as illustrated in figure 23. Hence the total concavity due to the join is $\alpha + \alpha' = (180^0 - \psi_1 - \psi_2 - \psi_3 - \psi_4)$. In order to establish a lower bound for the total concavity, we need to find upper bounds for the angles ψ_i $(i = 1$ to $4)$, and we can use the convexity of Σ_1 and Σ_2 to do this. Since the scaling functions of Σ_1 and Σ_2 are convex, the maximum possible value of ψ_2 is ϕ_2 (shown in figure 23), which is exactly $tan^{-1}w_2/2(l_2 - w_2 cot(\omega))$, and approximately $tan^{-1}(w_2/l_2)$. The maximum possible value of ψ_1 is ϕ_1, which is approximately $tan^{-1}(w_1/d_1)$ (see figure 23); and similarly for ψ_3 and ψ_4. These approximations hold provided that the cones are long relative to their widths, and d_1 is not near 0 or l_1 -- i.e. the join is not near either end of Σ_1.

Theorems 7 and 8 were stated precisely in the text and their proofs, which are straightforward, are omitted.